

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 12:59:40 ; Search time 55 Seconds
(without alignments)
570.232 Million cell updates/sec

Title: US-09-784-553C-19

Perfect score: 605
Sequence: 1 KPGRTVNLQYLKRVNKKAL.....IVLMAQTLEKIFLQKVASMP 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605	100.0	754	2	AAV07027
2	605	100.0	801	7	ADG35076
3	598	98.8	303	4	AB95802
4	542	89.6	140	4	AAU16619
5	542	89.6	140	6	ABU55688
6	542	89.6	235	4	AAU16206
7	542	89.6	235	6	ABU55275
8	500	82.6	688	3	AAV57898
9	500	82.6	1390	5	ABU65152
10	498	82.3	1937	4	AB958985
11	459	75.9	947	2	AAW81158
12	459	75.9	947	2	AAV07114
13	379	62.6	573	6	ABR41326
14	281.5	46.5	513	4	AB868341
15	269	44.5	330	6	ABR41589
16	237	39.2	648	4	ABR63109
17	221.5	36.6	703	5	AB93466
18	208	34.4	743	3	AGS52556
19	208	34.4	764	3	AGS52555
20	208	34.4	813	3	AGS52554
21	207	34.2	124	3	AAU32653
22	204	33.7	270	3	AGG42245
23	204	33.7	276	3	AGG42244
24	204	33.7	386	3	AGG42243
25	204	33.7	386	6	ABP70522

ALIGNMENTS

RESULT 1

AAV07027

ID AAV07027 standard; protein; 754 AA.

XX AC AAV07027;

DT 02-JUL-1999 (first entry)

DE Breast cancer associated antigen precursor sequence.

XX KW Cancer associated antigen; diagnosis; research; treatment; human; breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer; prostate cancer.

XX OS Homo sapiens.

XX PN WO9904265-A2.

XX PD 28-JAN-1999.

XX PF 15-JUL-1998; 98WO-US014679.

XX PR 17-JUL-1997; 97US-00896164.

XX PR 10-OCT-1997; 97US-0061599P.

XX PR 10-OCT-1997; 97US-0061765P.

XX PR 10-OCT-1997; 97US-00948705.

XX PR 11-OCT-1997; 97GB-00021697.

XX PR 22-JUN-1998; 98US-00102322.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;

XX O'hare M, Obata Y, Pfeundscht M, Tureci O, Sahin U;

XX WPI; 1999-132448/11.

XX New isolated cancer associated nucleic acids and polypeptides - isolated

XX using sera from cancer patients, used to develop products for the

XX diagnosis, monitoring or treatment of cancers.

XX Disclosure; Page 404-405; 787pp; English.

XX The invention relates to a method for diagnosing a disorder characterised by expression of a human cancer associated antigen precursor coded for by a nucleic acid molecule (NAM). The method comprises: (a) contacting a biological sample isolated from a subject with an agent that specifically binds to the NAM, an expression product or a fragment of an expression product complexed with an HLA molecule; and (b) determining the


```

SQ      Sequence 303 AA;
Query Match      98.8%; Score 598; DB 4; Length 303;
Best Local Similarity 99.1%; Pred. No. 6.7e-68;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 KGRVTNQLYVHKVVMKALWKHQFANFRQPDVAVKLGLEDYHKIKQPDMDGTIKRRL 60
      74 KGRVTNQLYVHKVVMKALWKHQFANFRQPDVAVKLGLEDYHKIKQPDMDGTIKRRL 133
Db
QY      61 ENNYWAAASECMQDNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
      134 ENNYWAAASECMQDNTMTFTNCYIYNKPTDDIVLMAQTPEKIFLOKVASMP 184
Db
RESULT 4
AAU16619
ID      AAU16619 standard; protein; 140 AA.
XX
AC      AAU16619;
DT
DT      07-NOV-2001 (first entry)
DE
DE      Human novel secreted protein, Seq ID 1572.
XX
KW      Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW      cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW      antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW      secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW      cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW      cerebral ischaemia; angiogenesis; nervous system disorder;
KW      Alzheimer's disease; infection; ocular disorder; corneal infection;
KW      wound healing; epithelial cell proliferation; skin ageing; food additive;
KW      preservative; antiproliferative.
XX
OS      Homo sapiens.
XX
PN      WO200155322-A2.
XX
PD      02-AUG-2001.
XX
FF      17-JAN-2001; 2001WO-US0001341.
XX
XX      31-JAN-2000; 2000US-0179065P.
PR      04-FEB-2000; 2000US-0180628P.
PR      24-FEB-2000; 2000US-0184664P.
PR      02-MAR-2000; 2000US-0186350P.
PR      16-MAR-2000; 2000US-0189874P.
PR      17-MAR-2000; 2000US-0190076P.
PR      18-APR-2000; 2000US-0198123P.
PR      19-MAY-2000; 2000US-0205515P.
PR      07-JUN-2000; 2000US-0209467P.
PR      28-JUN-2000; 2000US-0214886P.
PR      30-JUN-2000; 2000US-0215135P.
PR      07-JUL-2000; 2000US-0216847P.
PR      07-JUL-2000; 2000US-0216880P.
PR      11-JUL-2000; 2000US-0217487P.
PR      11-JUL-2000; 2000US-0217496P.
PR      14-JUL-2000; 2000US-0218290P.
PR      26-JUL-2000; 2000US-0220963P.
PR      26-JUL-2000; 2000US-0220964P.
PR      14-AUG-2000; 2000US-0224518P.
PR      14-AUG-2000; 2000US-0224519P.
PR      14-AUG-2000; 2000US-0225213P.
PR      14-AUG-2000; 2000US-0225214P.
PR      14-AUG-2000; 2000US-0225266P.
PR      14-AUG-2000; 2000US-0225267P.
PR      14-AUG-2000; 2000US-0225268P.
PR      14-AUG-2000; 2000US-0225270P.
PR      14-AUG-2000; 2000US-0225447P.
PR      14-AUG-2000; 2000US-0225575P.
PR      14-AUG-2000; 2000US-0225758P.
PR      14-AUG-2000; 2000US-0225759P.
PR      18-AUG-2000; 2000US-0226279P.
PR      22-AUG-2000; 2000US-0226861P.
PR      22-AUG-2000; 2000US-0226868P.
PR      22-AUG-2000; 2000US-0227182P.
PR      23-AUG-2000; 2000US-0227009P.
PR      30-AUG-2000; 2000US-0228924P.
PR      01-SEP-2000; 2000US-0229287P.
PR      01-SEP-2000; 2000US-0229343P.
PR      01-SEP-2000; 2000US-0229344P.
PR      01-SEP-2000; 2000US-0229345P.
PR      05-SEP-2000; 2000US-0229509P.
PR      05-SEP-2000; 2000US-0229513P.
PR      06-SEP-2000; 2000US-0230437P.
PR      06-SEP-2000; 2000US-0230438P.
PR      08-SEP-2000; 2000US-0231242P.
PR      08-SEP-2000; 2000US-0231243P.
PR      08-SEP-2000; 2000US-0231244P.
PR      08-SEP-2000; 2000US-0231413P.
PR      08-SEP-2000; 2000US-0231414P.
PR      08-SEP-2000; 2000US-0232080P.
PR      08-SEP-2000; 2000US-0232081P.
PR      12-SEP-2000; 2000US-0231968P.
PR      14-SEP-2000; 2000US-0232397P.
PR      14-SEP-2000; 2000US-0232398P.
PR      14-SEP-2000; 2000US-0232399P.
PR      14-SEP-2000; 2000US-0232400P.
PR      14-SEP-2000; 2000US-0232401P.
PR      14-SEP-2000; 2000US-0233063P.
PR      14-SEP-2000; 2000US-0233064P.
PR      14-SEP-2000; 2000US-0233065P.
PR      21-SEP-2000; 2000US-0234223P.
PR      21-SEP-2000; 2000US-0234274P.
PR      25-SEP-2000; 2000US-0234977P.
PR      25-SEP-2000; 2000US-0234988P.
PR      26-SEP-2000; 2000US-0235484P.
PR      27-SEP-2000; 2000US-0235834P.
PR      27-SEP-2000; 2000US-0235836P.
PR      29-SEP-2000; 2000US-0236367P.
PR      29-SEP-2000; 2000US-0236368P.
PR      29-SEP-2000; 2000US-0236369P.
PR      29-SEP-2000; 2000US-0236370P.
PR      02-OCT-2000; 2000US-0236802P.
PR      02-OCT-2000; 2000US-0237037P.
PR      02-OCT-2000; 2000US-0237038P.
PR      02-OCT-2000; 2000US-0237039P.
PR      02-OCT-2000; 2000US-0237040P.
PR      13-OCT-2000; 2000US-0239935P.
PR      13-OCT-2000; 2000US-0239937P.
PR      20-OCT-2000; 2000US-0240960P.
PR      20-OCT-2000; 2000US-0241231P.
PR      20-OCT-2000; 2000US-0241785P.
PR      20-OCT-2000; 2000US-0241786P.
PR      20-OCT-2000; 2000US-0241787P.
PR      20-OCT-2000; 2000US-0241808P.
PR      20-OCT-2000; 2000US-0241809P.
PR      01-NOV-2000; 2000US-0244617P.
PR      08-NOV-2000; 2000US-0246474P.
PR      08-NOV-2000; 2000US-0246475P.
PR      08-NOV-2000; 2000US-0246476P.
PR      08-NOV-2000; 2000US-0246477P.
PR      08-NOV-2000; 2000US-0246478P.
PR      08-NOV-2000; 2000US-0246523P.
PR      08-NOV-2000; 2000US-0246524P.
PR      08-NOV-2000; 2000US-0246525P.
PR      08-NOV-2000; 2000US-0246526P.
PR      08-NOV-2000; 2000US-0246527P.
PR      08-NOV-2000; 2000US-0246528P.
PR      08-NOV-2000; 2000US-0246532P.
PR      08-NOV-2000; 2000US-0246609P.
PR      08-NOV-2000; 2000US-0246610P.

```

09-NOV-2000; 2000US-0246611P.
 08-NOV-2000; 2000US-0246613P.
 17-NOV-2000; 2000US-0249207P.
 17-NOV-2000; 2000US-0249208P.
 17-NOV-2000; 2000US-0249209P.
 17-NOV-2000; 2000US-0249210P.
 17-NOV-2000; 2000US-0249211P.
 17-NOV-2000; 2000US-0249212P.
 17-NOV-2000; 2000US-0249213P.
 17-NOV-2000; 2000US-0249214P.
 17-NOV-2000; 2000US-0249215P.
 17-NOV-2000; 2000US-0249216P.
 17-NOV-2000; 2000US-0249217P.
 17-NOV-2000; 2000US-0249218P.
 17-NOV-2000; 2000US-0249244P.
 17-NOV-2000; 2000US-0249245P.
 17-NOV-2000; 2000US-0249264P.
 17-NOV-2000; 2000US-0249265P.
 17-NOV-2000; 2000US-0249297P.
 17-NOV-2000; 2000US-0249299P.
 17-NOV-2000; 2000US-0249300P.
 01-DEC-2000; 2000US-0250160P.
 05-DEC-2000; 2000US-0251030P.
 05-DEC-2000; 2000US-0251988P.
 06-DEC-2000; 2000US-0256713P.
 08-DEC-2000; 2000US-0251479P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251989P.
 08-DEC-2000; 2000US-0251990P.
 11-DEC-2000; 2000US-0254097P.
 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-488783/53.
 N-PSDB; AAS26606.

New nucleic acid molecules encoding 461 human secreted proteins for
 diagnosing, preventing, treating or ameliorating medical conditions and
 used as food additives or preservatives.

Claim 11; SEQ ID NO 1572; 980pp; English.

The invention relates to isolated nucleic acid molecules and their
 encoded secreted proteins. The nucleic acids and proteins are used to
 prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 in diagnosing a pathological condition or susceptibility to a
 pathological condition. Antibodies to the proteins can also be used in
 alleviating symptoms associated with the disorders and in diagnostic
 immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
 (ELISA). Disorders which are diagnosed or treated include autoimmune
 diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 nervous system disorders e.g. Alzheimer's disease, infections caused by
 bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 and many other disorders listed in the specification. The polypeptides
 can also be used to aid wound healing and epithelial cell proliferation,
 to prevent skin aging due to sunburn, to maintain organs before
 transplantation, for supporting cell culture of primary tissues, to
 regenerate tissues and in chemotaxis. The polypeptides can also be used
 as a food additive or preservative to increase or decrease storage
 capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 minerals, cofactors and other nutritional components, the present
 sequence represents a novel secreted protein of the invention. Note: The
 sequence data for this patent did not form part of the printed

Query Match 89.6%; Score 542; DB 4; Length 140;
 Best Local Similarity 87.4%; Pred. No. 3,7e-61;
 Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
 QY 1 KPGRTNQLQYVVKVVKLWKHQFAWPFQPDVAVKGLPDYHKIIFQPMDMGTIKRL 60
 Db 5 KPGRTNQLQYVVKVVKLWKHQFAWPFQPDVAVKGLPDYHKIIFQPMDMGTIKRL 64
 QY 61 ENNYWASACQDNTMTNCTIYKPTDDIVLVAQTLEKIFLOKVASMP 111
 Db 65 ENNYWASACQDNTMTNCTIYKPTDDIVLVAQTLEKIFLOKVASMP 115

RESULT 5

ABU55688
 ID ABU55688 standard; protein; 140 AA.

XX AC ABU55688;

XX DT 18-MAR-2003 (first entry)

XX DE Human novel polypeptide #775.

XX KW Human; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.

XX OS Homo sapiens.

XX PN US2002132753-A1.

XX PD 19-SEP-2002.

XX PF 17-JAN-2001; 2001US-00764864.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 14-AUG-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 22-AUG-2000; 2000US-0226868P.

XX PR 30-AUG-2000; 2000US-0228924P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229344P.

XX PR 01-SEP-2000; 2000US-0229345P.

XX PR 05-SEP-2000; 2000US-0229509P.

XX PR 05-SEP-2000; 2000US-0229513P.

XX PR 21-SEP-2000; 2000US-0231413P.

XX PR 21-SEP-2000; 2000US-0234223P.

XX PR 25-SEP-2000; 2000US-0234997P.

XX PR 27-SEP-2000; 2000US-0235834P.

XX PR 29-SEP-2000; 2000US-0236327P.

XX PR 29-SEP-2000; 2000US-0236367P.

XX PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR N-PSDB; ABX73947.
DR
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX Claim 11; SEQ ID NO 1572; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
XX Sequence 140 AA;
XX
Query Match 89.6%; Score 542; DB 6; Length 140;
Best Local Similarity 87.4%; Pred. No. 3.7e-61;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Oy 1 KPGKRVNQLQYLHKVYKMLWKEQFAPRPQVDVAVKIGLPDYVHKTIKOPMDMGITKRL 60
Db 5 KPGKRVNQLQYQNVVVKLWKEQFAPRPQVDVAVKIGLPDYVHKTIKOPMDMGITKRL 64
Oy 61 ENNYWAASECMDQFNTMTNCYIYNKPTDDIVLMAQTLEKIFLQKVASP 111
Db 65 ENNYWAASECMDQFNTMTNCYIYNKPTDDIVLMAQTLEKIFLQKVASP 115
RESULT 6
AAU16206
ID AAU16206 standard; protein: 235 AA.
XX
XX AAU16206;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human novel secreted protein, Seq ID 1159.
XX
XX
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
XX Homo sapiens.
OS
XX WO200155322-A2.
PN
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001341.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226888P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.

DR WPI; 2000-072605/06.
 DR N-PSDB; AA256719.
 XX
 PT Proteins, polynucleotides, vectors, host cells and antibodies used to
 PT diagnose, treat or prevent immune, reproductive, smooth muscle,
 PT neurological, gastrointestinal, developmental and cell proliferative
 PT disorders.
 XX
 PS Claim 1; Page 128-130; 229pp; English.
 XX
 CC AA256698 to AA256776 encode AAY57877 to AAY57955 which represent human
 CC transmembrane proteins designated HTMPN-1 to HTMPN-79, respectively. The
 CC transmembrane protein have immunospecific, antiproliferative and
 CC neuroprotective activities. The human transmembrane proteins,
 CC polynucleotides encoding them and other compositions and methods from the
 CC present invention, can be used for the diagnosis, treatment or prevention
 CC of immune, reproductive, smooth muscle, neurological, gastrointestinal,
 CC developmental and cell proliferative disorders. The HTMPN's can be used
 CC to treat or prevent disorders associated with a decreased expression or
 CC activity of HTMPN
 CC
 SQ Sequence 688 AA;
 Query Match 82.6%; Score 500; DB 3; Length 688;
 Best Local Similarity 80.2%; Pred. No. 7.6e-55;
 Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
 QY 1 KPGRTNQLQVHLKVMKALWKHOFAPFPQVDAVKGLPDYHKIKQPMDMGTIKRRL 60
 DB 55 KPGRTNQLQVHLKVMKALWKHOFAPFPQVDAVKGLPDYHKIKQPMDMGTIKRRL 114
 QY 61 ENNYWAAECIQDFTNFTNCYINXPTDDIVLMAOTLEKIFLOKVASMP 111
 DB 115 ENNYWAAECIQDFTNFTNCYINXPTDDIVLMAOTLEKIFLOKVASMP 165
 RESULT 9
 ABU65152
 ID ABU65152 standard; protein; 1390 AA.
 XX
 AC ABU65152;
 XX
 DT 20-MAY-2003 (first entry)
 XX
 DE Human NOV79a protein.
 XX
 KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO200272757-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US006908.
 XX
 PR 08-MAR-2001; 2001US-0274101P.
 PR 08-MAR-2001; 2001US-0274194P.
 PR 08-MAR-2001; 2001US-0274281P.
 PR 08-MAR-2001; 2001US-0274322P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.
 PR 13-MAR-2001; 2001US-0275578P.
 PR 13-MAR-2001; 2001US-0275579P.
 PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 16-MAR-2001; 2001US-0276776P.
 PR 19-MAR-2001; 2001US-0276994P.
 PR 20-MAR-2001; 2001US-0277239P.
 PR 20-MAR-2001; 2001US-0277321P.
 PR 20-MAR-2001; 2001US-0277327P.

PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0278999P.
 PR 28-MAR-2001; 2001US-0279036P.
 PR 30-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0277733P.
 PR 30-MAR-2001; 2001US-0279955P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 04-APR-2001; 2001US-0280900P.
 PR 13-APR-2001; 2001US-0281194P.
 PR 30-APR-2001; 2001US-0283675P.
 PR 02-MAY-2001; 2001US-0287424P.
 PR 03-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 15-MAY-2001; 2001US-0288528P.
 PR 16-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318776P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 14-NOV-2001; 2001US-0333272P.
 PR 21-NOV-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 03-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 07-MAR-2002; 2002US-00092900.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
 PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Pattursajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
 PI Fernandes ER, Casman SU, Malyankar UM, Gerlach V, Liu Y, Anderson D;
 PI Spaderina SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
 PI Lepley DM, Rieger DK;
 XX
 XX WPI; 2002-723332/78.
 DR N-PSDB; ABX97119.
 XX
 PT NOVX polypeptides and polynucleotides, useful for preventing or treating
 PT a disorder associated with aberrant NOVX expression or activity e.g.,
 PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 PT asthma.
 XX
 PS Claim 1; Page 323; 1103pp; English.
 XX
 CC This invention describes novel human NOVX polypeptides which have
 CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
 CC activity. Pharmaceutical compositions comprising the NOVX proteins or
 CC nucleic acid molecules or NOVX antibodies are useful for preventing or
 CC treating a disorder associated with aberrant NOVX expression or activity

CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABUS5041-ABU65218 represent the NOVX polypeptides encoded by
CC ABX97008-ABX97185
XX
SQ Sequence 1390 AA;

Query Match 82.6%; Score 500; DB 5; Length 1390;
Best Local Similarity 80.2%; Pred. No. 1.9e-54;
Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHLKVKWALWKHOFAPWPPQPDVAVKLGDPYHKIIPQMDMGTIKRL 60
DB 55 KPGRTVNLQYLHLKVKWALWKHOFAPWPPQPDVAVKLGDPYHKIIPQMDMGTIKRL 114
QY 61 ENNYWAASECQDFTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 115 ENNYWAASECQDFTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 165

RESULT 10
ABB58985
ID ABB58985 standard; protein; 1937 AA.
XX
AC ABB58985;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 3747.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
FR 11-JUL-2000; 2000US-00614150.
XX
PA (PEXE) PE CORP NY.
XX
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL03088.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 3747; 21pp + Sequence Listing; English.
XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB589737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1937 AA;

Query Match 82.3%; Score 498; DB 4; Length 1937;
Best Local Similarity 77.5%; Pred. No. 5.4e-54;
Matches 86; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHLKVKWALWKHOFAPWPPQPDVAVKLGDPYHKIIPQMDMGTIKRL 60
DB 31 RPSGNTNLQQLYLLTVMKVIMKHHSWPFQFPVDKAKLNLDPYHKIIPQMDMGTIKRL 90
QY 61 ENNYWAASECQDFTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 91 ENNYWAASECQDFTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 141

RESULT 11
AAW81168
ID AAW81168 standard; protein; 947 AA.
XX
AC AAW81168;
XX
DT 05-MAR-1999 (first entry)
XX
DE Transcriptional regulatory factor RING3.
XX
KW Human; transcriptional regulatory factor; RING3; TSB; cancer;
KW testis specific; bromodomain; testicular cell proliferation.
XX
OS Homo sapiens.
XX
PN WO9848015-A1.
XX
PD 29-OCT-1998.
XX
PF 17-APR-1998; S8WO-JP001782.
XX
PR 18-APR-1997; 97JP-00116402.
XX
PA (CHUS) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Jones MH;
XX
DR WPI; 1998-583658/49.
DR N-PSDB; AAV68343.
XX
PT Transcriptional regulator gene containing bromodomain sequence - may be
PT expressed in testis tissue and is useful in treatment of cancer and other
PT proliferative disorders.
XX
PS Claim 1; Page 19-24; 42pp; Japanese.
XX
XX
CC The present sequence represents the human transcriptional regulatory
CC factor RING3, which is isolated from testicular cells. RING3 contains a
CC testis specific bromodomain (TSB) which is expressed specifically in
CC testis tissue and also expressed in certain tumour lines. The transgenic
CC cells may be used to express RING3 which is a TSB expression protein. The
CC TSB expression product can be used in the treatment of cancer and other
CC proliferative disorders, and in screening of compounds for ability to
CC bind to it (e.g. for use as drugs by modulation of transcriptional
CC regulation). DNA capable of hybridising to RING3 polynucleotides may be
CC used for construction of probes and primers
XX
SQ Sequence 947 AA;

Query Match 75.9%; Score 459; DB 2; Length 947;
Best Local Similarity 73.0%; Pred. No. 2.1e-49;
Matches 81; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHLKVKWALWKHOFAPWPPQPDVAVKLGDPYHKIIPQMDMGTIKRL 60
DB 24 KNGRITNLQYLQKVLKDLWKHSFWSWPFQFPVDKAKLNLDPYHKIIPQMDMGTIKRL 83
QY 61 ENNYWAASECQDFTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 84 ENKYAKASECIEDFTMTNFTNFSNCYLYNKPDDIVLMAQALEKLFNQKLSOMP 134

RESULT 12

CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a DITHP protein which has intracellular
CC signalling activity. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 573 AA;

Query Match 62.6%; Score 379; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 2e-39;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 DYHKIIKQPMGCTIKRLENNYWAASCMQDFTMTNCTIYNNKPTDDIVLMAQTLEK 101
DB 5 DYHKIIKQPMGCTIKRLENNYWAASCMQDFTMTNCTIYNNKPTDDIVLMAQTLEK 64
QY 102 IFLOKVASMP 111
DB 65 IFLOKVASMP 74

RESULT 14
ABB68341
ID ABB68341 standard; protein; 513 AA.
AC ABB68341;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 31815.

XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.

XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEXE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL12444.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX Disclosure; SEQ ID NO 31815; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
XX sequences (ABU1840-ABU16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 513 AA;

Query Match 46.5%; Score 281.5; DB 4; Length 513;
Best Local Similarity 44.5%; Pred. No. 5.8e-27;
Matches 49; Conservative 25; Mismatches 35; Indels 1; Gaps 1;
QY 1 KPGRTVITQLQYLHKVNMKALWKHQFAWFFRQPVDAVKGLPLDYHKIIKQPMGCTIKRRL 60
DB 36 RPRGRTNIBEL-KSVLNCLEWRNRFVYHFRHVDVSLGVDFYHVVKHPMDLSTIKRL 94
QY 61 ENNYWAASCMQDFTMTNCTIYNNKPTDDIVLMAQTLEKIFLOKVASM 110
DB 95 HNKYYWQASEALEDFKLIFDNCLLYNLEGSFVYQAGKLLMEAFYMRMESI 144

RESULT 15
ABR41589
ID ABR41589 standard; protein; 330 AA.
XX
AC ABR41589;
XX
DT 02-JUN-2003 (first entry)
XX
DE Human DITHP antigen recognition protein.

XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
XX cancer; cell proliferative disorder; autoimmune disorder;
XX inflammatory disorder; infection; hormonal disorder; metabolic disorder;
XX neurological disorder; gastrointestinal disorder; transport disorder;
XX connective tissue disorder; drug screening; proteome analysis;
XX gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
XX disease model; toxicological testing; transcript imaging;
XX antigen recognition.

XX Homo sapiens.
XX WO200297031-A2.
XX
PD 05-DEC-2002.

XX 27-MAR-2002; 2002WO-US010056.
XX
XX 28-MAR-2001; 2001US-0279619P.
XX 29-MAR-2001; 2001US-0280067P.
XX 29-MAR-2001; 2001US-0280068P.
XX 16-MAY-2001; 2001US-0291280P.
XX 17-MAY-2001; 2001US-0291829P.
XX 17-MAY-2001; 2001US-0291849P.
XX 19-JUN-2001; 2001US-0299428P.
XX 20-JUN-2001; 2001US-0299776P.
XX 20-JUN-2001; 2001US-0300001P.
XX (INCY-) INCYTE GENOMICS INC.

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
XX Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
XX Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
XX Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-129518/12.
XX N-PSDB; ACC46527.

XX Novel human diagnostic and therapeutic polypeptide useful for identifying
XX test compound which specifically binds to a polypeptide encoded by human
XX diagnostic and therapeutic polynucleotide, and to induce antibodies.

XX Claim 27; SEQ ID NO 1124; 591pp; English.
XX The invention relates to novel human diagnostic and therapeutic
XX polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
XX proteins (DITHP; ABR41136-ABR41812). The invention also relates to
XX polynucleotide sequences at least 90% identical to the dithp cDNA
XX sequences of the invention; recombinant vectors, host cells and

transgenic organisms comprising a dithp nucleic acid sequence; the recombinant production of DITHP proteins; antibodies specific for DITHP proteins; microarrays comprising dithp nucleic acid sequences; methods of detecting dithp nucleotide and protein sequences; methods of screening for compounds which specifically bind a DITHP protein; and methods of assessing the toxicity of test compounds using a dithp hybridisation probe. Dithp nucleic acid sequences and DITHP proteins may be used in the diagnosis of a wide variety of conditions including cancer and other cell proliferative disorders; autoimmune or inflammatory disorders; bacterial, viral, fungal or parasitic infections; hormonal disorders; metabolic disorders; neurological disorders; gastrointestinal disorders; transport disorders; and connective tissue disorders. They may also be used to screen for modulators of protein activity or gene expression. DITHP proteins can additionally be used in analysis of the proteome of a tissue or cell type and to induce antibodies. The dithp nucleic acids are additionally useful in somatic or germline gene therapy of the disorders mentioned above, as a source of antisense sequences, as a source of probes and primers, in genotyping and identification of individuals, in the generation of transgenic animal models of human disease or knock in humanised animals, in toxicological testing, and in transcript imaging. The present sequence represents a DITHP protein which has antigen recognition activity. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

XX SQ Sequence 330 AA;

Query Match	44.5%;	Score 269;	DB 6;	Length 330;
Best Local Similarity	46.0%;	Pred. No. 1.3e-25;		
Matches	52;	Conservative	21;	Mismatches 38; Indels 2; Gaps 1;

QY	1	KPCRVTNQLVHLKVKVWKAHWQ--FAWPRQPDVAVKGLGDPYHKIIKQPDMDMTIKR	58
Db	153	KGKLSLSEQLKHCNGILKELLSKGAAYAWPFYKPDVAVKGLGDPYHKIIKQPDMDMTIKR	212
QY	59	RLNNYVWAASECQDFNTMFTNCYVYKPTDDIVLMAQTLEKIFLOKVASMP	111
Db	213	KVENRDYDAQFAADVRLMESNCYKYNPDHVVVMAVKLQVFFRYAMP	265

Search completed: July 9, 2004, 13:06:14
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:04:05 ; Search time 21 Seconds
(without alignments)
508.441 Million cell updates/sec

Title: US-09-784-553c-19
Perfect score: 605
Sequence: 1 KPGRTNQLQYLHKVWKAL.....IVLMAQTLEKIFLQKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	605	100.0	754	2 A56619	female sterile hom
2	594	98.2	733	2 T28145	RING3 kinase - chi
3	498	82.3	2038	2 A43742	female sterile hom
4	296	48.9	1250	2 T22845	hypothetical prote
5	279	46.1	1087	2 T22847	hypothetical prote
6	221.5	36.6	703	2 T48500	kinase-like protei
7	215	35.5	578	2 T40984	transcription fact
8	204	33.7	400	2 T00472	probable RING3 pro
9	201	33.2	461	2 D96757	hypothetical prote
10	198	32.7	832	1 S71788	P/CAR protein - hu
11	195.5	32.3	374	2 T33328	hypothetical prote
12	195	32.2	678	2 T49984	bromodomain protei
13	190.5	31.5	361	2 T42517	bromodomain protei
14	188	31.1	766	2 A86198	hypothetical prote
15	187	30.9	440	2 H86312	F2HUS.2 protein -
16	184.5	30.5	2441	2 S39161	CREB-binding prote
17	183.5	30.3	638	2 S67605	hypothetical prote
18	182	30.1	439	1 S28051	transcription fact
19	181.5	30.0	2414	2 A54277	transcription adap
20	180.5	29.8	2440	2 S39162	transcription coac
21	179.5	29.7	586	2 T47620	histon acetyltrans
22	176	29.1	405	2 T21433	hypothetical prote
23	176	29.1	432	2 T21435	hypothetical prote
24	176	29.1	510	2 T21430	hypothetical prote
25	176	29.1	686	2 S55955	bromodomain protei
26	172	28.4	454	2 T37933	transcription acti
27	168.5	27.9	759	2 E96613	hypothetical prote
28	165	27.4	3190	2 T13828	CREB-binding prote
29	162	26.8	2027	2 S60323	hypothetical prote

30 162 26.8 2056 2 G89564 protein R10E11.1 [

31 160.5 26.5 369 2 T46098 hypothetical prote

32 160.5 26.5 1051 2 S55259 TIF1 protein - mou

33 158 26.1 476 1 S71789 GCN5 protein - hum

34 157 26.0 757 2 S68142 probable transcrip

35 149 24.6 449 2 T12495 hypothetical prote

36 137.5 22.7 979 2 T40006 hypothetical prote

37 136.5 22.6 811 2 T08738 hypothetical prote

38 136 22.5 1893 1 A40262 transcription init

39 135 22.3 1680 2 T41628 probable transcrip

40 134.5 22.2 1479 2 T17401 transcripition regu

41 128 21.2 1865 1 T48155 transcripition init

42 127.5 21.1 2068 2 A47371 transcripition init

43 127 21.0 1633 2 JC5056 polybromo 1 - chic

44 127 21.0 1638 2 A42091 transcripition acti

45 127 21.0 1879 2 T19481 hypothetical prote

ALIGNMENTS

RESULT 1

A56619 female sterile homeotic (fsh) homolog RING3 - human

C:Species: Homo sapiens (man)

C>Date: 21-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 20-Sep-1999

C:Accession: A56619; S18860; S40781

R:Beck, S.; Hanson, I.; Kelly, A.; Pappin, D.J.; Trowsdale, J.

DNA Seq. 2, 203-210, 1992

A:Title: A homologue of the Drosophila female sterile homeotic (fsh) gene in the class I)

A:Reference number: A56619; MUID:92329974; PMID:1352711

A:Accession: A56619

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-754 <BEC>

A:Cross-references: EMBL:X62083; NID:g31471; PIDN:AAA68890.1; PID:g31472; EMBL:M80613; NJ

A>Note: sequence extracted from NCBI backbone (NCBIP:108781)

C:Genetics:

A:Gene: RING3

C:Superfamily: unassigned bromodomain proteins; bromodomain homology

C:Keywords: duplication

F:52-109/Domain: bromodomain homology <BRO1>

F:325-382/Domain: bromodomain homology <BRO2>

Query Match 100.0%; Score 605; DB 2; Length 754;

Best Local Similarity 100.0%; Pred. No. 7.6e-58;

Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVWKALWKHQFAWPFROPVDAVKGLPDYHKIHKQPMDMGTIKERL 60

Db 24 KPGRTNQLQYLHKVWKALWKHQFAWPFROPVDAVKGLPDYHKIHKQPMDMGTIKERL 83

QY 61 ENNYWAASECQDFTNTMFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111

Db 84 ENNYWAASECQDFTNTMFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 134

RESULT 2

RING3 kinase - chicken

C:Species: Gallus Gallus (chicken)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28145

R:Milne, S.; Kaufman, J.; Beck, S.

submitted to the EMBL Data Library, May 1998

A:Description: DNA sequencing and analysis of the chicken major histocompatibility comple

A:Reference number: Z20475

A:Accession: T28145

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-733 <MIL>

A:Cross-references: EMBL:AL023516; NID:e1292539; PID:e1292549; PIDN:CAA18965.1

A:Experimental source: clone CB12

```
C:Genetics:
A:Gene: RING3
A:Map position: 16
A:Introns: 64/3; 110/3; 158/1; 227/3; 351/3; 394/3; 479/3; 546/2; 650/1; 691/1
A:Superfamily: unassigned bromodomain proteins; bromodomain homology
F:52-109/Domain: bromodomain homology <BRO1>
F:323-380/Domain: bromodomain homology <BRO2>

Query Match      98.2%; Score 594; DB 2; Length 733;
Best Local Similarity 97.3%; Pred. No. 1.2e-56;
Matches 108; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHGFANFPFQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 60
DB 24 KPGRTVNLQYLHKVVMKALWKHGFANFPFQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 83

QY 61 ENNYWAAECMQDFTNTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 84 ENNYWGAECMQDFTNTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 134

RESULT 3
A43742
female sterile homeotic protein, 205K - fruit fly (Drosophila melanogaster)
N:Alternate names: membrane protein fish, 205K
N:Contains: female sterile homeotic protein, 110K
C:Species: Drosophila melanogaster
C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 20-Sep-1999
C:Accession: A43742; B43742
R:Haynes, S.R.; Mazer, B.A.; Bhatia-Dey, N.; Dawid, I.B.
Dev. Biol. 134, 246-257, 1989
A:Title: The Drosophila fish locus, a maternal effect homeotic gene, encodes apparent mem
A:Reference number: A43742; MUID:89276730; PMID:2567251
A:Accession: A43742
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2038 <HAY>
A:Cross-references: EMBL:M22221; NID:g157452; PIDN:AAA28540.1; PID:g157453
A:Accession: B43742
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1106 <HAY>
A:Cross-references: EMBL:M22222
C:Genetics:
A:Gene: fish
A:Cross-references: FlyBase:FBgn0004656
C:Superfamily: unassigned bromodomain proteins; bromodomain homology
C:Keywords: alternative splicing; transmembrane protein
F:1-2038/Product: female sterile homeotic protein, 205K #status predicted <HAY>
F:1-1106/Product: female sterile homeotic protein, 110K #status predicted <HAY>
F:59-116/Domain: bromodomain homology <BRO1>
F:503-560/Domain: bromodomain homology <BRO2>

Query Match      82.3%; Score 498; DB 2; Length 2038;
Best Local Similarity 77.5%; Pred. No. 1.2e-45;
Matches 86; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHGFANFPFQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 60
DB 31 RPGRTVNLQYLHKVVMKALWKHGFANFPFQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 90

QY 61 ENNYWAAECMQDFTNTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 91 ENNYWAAKETIQDFTNTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 141

RESULT 4
T22845
hypothetical protein F57C7.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
C:Accession: T22845
R:White, S.

submitted to the EMBL Data Library, February 1996
A:Reference number: Z19625
A:Accession: T22845
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1250 <WIL>
A:Cross-references: EMBL:Z69646; PIDN:CAA93473.1; GSPDB:GNO0028; CESP:F57C7.1a
A:Experimental source: clone F57C7
C:Genetics:
A:Gene: CESP:F57C7.1a
A:Map position: X
A:Superfamily: bromodomain homology
F:307-364/Domain: bromodomain homology <BRO1>
F:619-676/Domain: bromodomain homology <BRO2>

Query Match      48.9%; Score 296; DB 2; Length 1250;
Best Local Similarity 46.8%; Pred. No. 7e-24;
Matches 52; Conservative 23; Mismatches 36; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHGFANFPFQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 60
DB 279 KPTRHTNCLDFVLFVTVKDALKHSHWPFQLPVDAIKLEIPEYHNI VNTPMDLSTIKRRL 338

QY 61 ENNYWAAECMQDFTNTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 339 RNLVWCAEDAIKDINQVFINCYSPENPEYDVYKMAKLEKQVLSQLTQLP 399

RESULT 5
T22847
hypothetical protein F57C7.1b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
C:Accession: T22847
R:White, S.

submitted to the EMBL Data Library, February 1996
A:Reference number: Z19625
A:Accession: T22847
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1087 <WIL>
A:Cross-references: EMBL:Z69646; PIDN:CAA93475.1; GSPDB:GNO0028; CESP:F57C7.1b
A:Experimental source: clone F57C7
C:Genetics:
A:Gene: CESP:F57C7.1b
A:Map position: X
A:Introns: 262/3; 351/3; 391/3; 627/3; 795/2; 957/1
C:Superfamily: bromodomain homology
F:307-364/Domain: bromodomain homology <BRO1>
F:579-636/Domain: bromodomain homology <BRO2>

Query Match      46.1%; Score 279; DB 2; Length 1087;
Best Local Similarity 44.1%; Pred. No. 4.3e-22;
Matches 49; Conservative 24; Mismatches 38; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHGFANFPFQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 60
DB 279 KPTRHTNCLDFVLFVTVKDALKHSHWPFQLPVDAIKLEIPEYHNI VNTPMDLSTIKRRL 338

QY 61 ENNYWAAECMQDFTNTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 339 RNLVWCAEDAIKDINTLFDNCKFNDRNDIYMCENIEGVVGQGLEWMP 399

RESULT 6
T48600
kinase-like protein - Arabidopsis thaliana
N:Alternate names: protein F18022.60
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000
C:Accession: T48600
R:Sevan, M.; Murphy, G.; Ridley, G.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lea
```

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493

A:Accession: T48600

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-703 <BEV>

A:Cross-references: EMBL:DB

A:Experimental source: cultivar Columbia; BAC clone F18022

C:Genetics:

A:Map position: 5

A:Introns: 370/3; 387/1; 423/1; 461/3; 484/1; 511/1; 549/3; 572/1; 619/3

A:Note: F18022.60

C:Superfamily: bromodomain homology

F:157-214/Domain: bromodomain homology <BRO>

Query Match 36.6%; Score 221.5; DB 2; Length 703;
Best Local Similarity 41.5%; Pred. No. 4.7e-16;
Matches 44; Conservative 15; Mismatches 44; Indels 3; Gaps 1;

QY 1 KGRVTNQLQYLHK---VVKALWKHQFAMPPRPQVDVAVKGLPDYHKIIKQPMDMGTHK 57
DB 126 KPTPTVAVMLKMQCEALLKRLMSHQYGVFNTFVVKLNILDFNVNIEHPMDLGTVK 185

QY 58 RLENNYVWAAECMDQDFNTMTNCYVYKPKTDDIVLMAQTLEKIF 103

DB 186 NKLTSCTYSCPSEFAADVRLTFSNANTYPPGNDVYVWADTLRKTF 231

RESULT 7

T40984 transcription factor baf1 homolog SPC1450.02 - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: protein SPC191.13

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 08-Dec-2000

C:Accession: T40984; T41225

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.

submitted to the EMBL Data Library, March 1999

A:Reference number: Z21962

A:Accession: T40984

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-578 <LYN1>

A:Cross-references: EMBL:AL049559; NID:94581510; PIDN:CAB40169.1; PID:g4581512; GSPDB:GN

A:Experimental source: strain 972h-; cosmid c1450

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.

submitted to the EMBL Data Library, March 1998

A:Reference number: Z21904

A:Accession: T41225

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-111 <LYN2>

A:Cross-references: EMBL:AL049644; NID:94678667; PIDN:CAB41059.1; PID:g4678679; GSPDB:GN

A:Experimental source: strain 972h-; cosmid c191

C:Genetics: <C145>

A:Gene: SPDB:SPCC1450.02

A:Map position: 3

C:Genetics: <C191>

A:Gene: SPDB:SPCC191.13

A:Map position: 1

C:Superfamily: bromodomain homology

F:109-166/Domain: bromodomain homology <BRO1>

F:279-336/Domain: bromodomain homology <BRO2>

Query Match 35.5%; Score 215; DB 2; Length 578;
Best Local Similarity 40.9%; Pred. No. 1.9e-15;
Matches 47; Conservative 15; Mismatches 49; Indels 4; Gaps 2;

QY 1 KGRVTNQLQY-LHKVVMKALWKHQ---FAMPPRPQVDVAVKGLPDYHKIIKQPMDMGTH 56
DB 247 KPRRRKNSQMRFCSTVLKELYKQVSEFAFPYQFVDPVACDPCDPYFDVIEKPMDLSTI 306

QY 57 KRLNNYVWAAECMDQDFNTMTNCYVYKPKTDDIVLMAQTLEKIFLOKVASMP 111

Db 307 QSKLNKNEYSTLEEFESDILLMFNNCTYNPPGTPVHVMGRLQENVFKEKWEARP 361

RESULT 8

T00472

probable RING3 protein [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein F1913.13

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001

C:Accession: T00472; C84762

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,

submitted to the EMBL Data Library, April 1998

C:Superfamily: Arabidopsis thaliana chromosome II BAC F1913 genomic sequence.

A:Reference number: Z14160

A:Accession: T00472

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-400 <ROU>

A:Cross-references: EMBL:AC004238; NID:g3033373; PID:g3033386

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; A

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: C84762

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <STO>

A:Cross-references: GB:AE002093; NID:g3033386; PIDN:AAAC12830.1; GSPDB:GN00139

C:Genetics:

A:Gene: Atg334900; F1913.13

A:Map position: 2

A:Introns: 6/3; 49/3; 120/3; 146/3; 218/3; 251/3; 271/2; 335/3

C:Superfamily: bromodomain homology

F:130-187/Domain: bromodomain homology <BRO>

Query Match 33.7%; Score 204; DB 2; Length 400;
Best Local Similarity 39.6%; Pred. No. 2e-14;
Matches 36; Conservative 26; Mismatches 29; Indels 0; Gaps 0;

QY 16 VNKALWKHQFAMPPRPQVDVAVKGLPDYHKIIKQPMDMGTHKRLNNYVWAAECMDQF 75

DB 117 MFRQIAQHKWAPFEPVDVKGLGHDYKVKYIEKPMDLGTHKQVSESYNVEIYADV 176

QY 76 NTMTNCYVYKPKTDDIVLMAQTLEKIFLOK 106

DB 177 RLVPKQAMRYNEKEDEVVWMAESLTERPEEK 207

RESULT 9

D36757

hypothetical protein T18X17.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96757

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.D.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96757

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-461 <STO>

A:Cross-references: GB:AE005173; NID:96598866; PIDN:AAF18720.1; GSPDB:GN00141

```

C:Superfamily: bromodomain homology
F:141-198/Domain: bromodomain homology <BRO>

Query Match          32.3%; Score 195.5; DB 2; Length 374;
Best Local Similarity 40.0%; Pred. No. 1.5e-13;
Matches 42; Conservative 12; Mismatches 42; Indels 9; Gaps 1;

QY 11 YLHKVWKAL-----WKHOFAMPFRQPDVAVKLGLPDYHKLIKQPMDMGTIKRRL 61
Db 114 HLHDLAKCLISLKFEFKSTHDSFTFFRKPVDVVLGLTDYHEVIKKPMDMTIRKKLI 173
QY 62 NNYWYAASECMQDNTFTNCYIYNKPTDDIVLMAQTLEKIFLQK 106
Db 174 GEEDYTAVEFKEDPKGLMINNCLTYNNEGDPVADFALQFRKKFAAK 218

RESULT 12
T49984
bromodomain protein-like - Arabidopsis thaliana
N:Alternate names: protein F12B17.100
C:Species: Schizosaccharomyces pombe
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
C:Accession: T49984
R:Bevan, M.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25026
A:Accession: T49984
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-878 <REV>
A:Cross-references: EMBL:AL353995; GSPDB:GN00063; ATSP:F12B17.100
A:Experimental source: cultivar Columbia; BAC clone F12B17
C:Genetics:
A:Gene: ATSP:F12B17.100
A:Map position: 5
A:Introns: 73/3; 560/3; 629/1; 656/2
C:Superfamily: bromodomain homology
F:269-326/Domain: bromodomain homology <BRO>

Query Match          32.2%; Score 195; DB 2; Length 678;
Best Local Similarity 37.4%; Pred. No. 3.5e-13;
Matches 40; Conservative 17; Mismatches 38; Indels 12; Gaps 1;

QY 16 VMKALWKHPAMPFRQPDVAVKLGLPDYHKLIKQPMDMGTIKRRLNNYWAASECMQDF 75
Db 256 ILVKMLKHKSWVFLNPVDVVGILGHYHRIVDKPMDLGTQVKNLEKGLVRSPIDFASDV 315
QY 76 NTWFTNCYIYNKPTDDIVLMAQ-----TLEKIFQKVASM 110
Db 316 RLFTNAMSINPKQDQVYLMAEKLISQFDVWNPNTLKRFEAQEVKVM 362

RESULT 13
T42517
bromodomain protein - fission yeast (Schizosaccharomyces pombe) (fragment)
C:Species: Schizosaccharomyces pombe
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C:Accession: T42517
R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs
A:Reference number: Z17323; MUID:98162722; PMID:9501991
A:Accession: T42517
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-361 <YOS>
A:Cross-references: EMBL:D89157; NID:gl749521; PIDN:BAA13819.1; PID:gl749522
A:Experimental source: strain PR745
C:Superfamily: bromodomain homology
F:164-121/Domain: bromodomain homology <BRO>

Query Match          31.5%; Score 190.5; DB 2; Length 361;
Best Local Similarity 40.4%; Pred. No. 5.2e-13;

```


Matches 40; Conservative 17; Mismatches 39; Indels 3; Gaps 1;
QY 16 VMKALWKHC--FANPFRQPVDAVKLGDPYHKIKQPMDMGTIKRRLNNYVWAAASECM 72
Db 48 VIKELKKQHEAYAFYFKVNPVTACGCPDYFKVKGHPMDLGTMONKLNHNEYASMKAFE 107
QY 73 QDFNTMFTNCYIYNKFTDDIVLMAQTLEKIFLQKVASMP 111
Db 108 ADWLMFMKCYKFNAGTGVHLMGKLESIFQKLWANKP 146
RESULT 14
A86198
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86198
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-766 <STO>
A:Cross-references: GB:AE005172; NID:g9844128; PIDN:AAF60220.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 31.1%; Score 188; DB 2; Length 766;
Best Local Similarity 43.0%; Pred. No. 2.4e-12;
Matches 40; Conservative 13; Mismatches 40; Indels 0; Gaps 0;
QY 16 VMKALWKHC--FANPFRQPVDAVKLGDPYHKIKQPMDMGTIKRRLNNYVWAAASECMQDF 75
Db 428 LLERLMKHKHGWVFNAPVDVKGGLLDYVYTIIEHPMDLGTIKSALMKNLYKSPREFAEADV 487
QY 76 NTMFTNCYIYNKFTDDIVLMAQTLEKIFLQKVA 108
Db 488 RLTFHNTYINPEGQDVHLMVATLLQIFERWA 520

RESULT 15
H86312
P2H15.2 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C:Accession: H86312
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H86312
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-440 <STO>
A:Cross-references: GB:AE005172; NID:g9665057; PIDN:AAF97259.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 30.9%; Score 187; DB 2; Length 440;
Best Local Similarity 38.9%; Pred. No. 1.6e-12;
Matches 37; Conservative 17; Mismatches 41; Indels 0; Gaps 0;
QY 16 VMKALWKHC--FANPFRQPVDAVKLGDPYHKIKQPMDMGTIKRRLNNYVWAAASECMQDF 75
Db 139 LLTKLMKHKSAWVFNVPVDKGLGHDYHNIVKPEMDLGTIKLGLKSLYKSPDFAEDV 198
QY 76 NTMFTNCYIYNKFTDDIVLMAQTLEKIFLQKVASM 110
Db 199 RLTFHNTYINPEGQDVHLMVATLLQIFERW 233

Search completed: July 9, 2004, 13:08:07
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:00:15 ; Search time 17 Seconds
(without alignments)
339.987 Million cell updates/sec

Title: US-09-784-553C-19

Perfect score: 605
Sequence: 1 KGRVTNQLYLHKVMKAL.....IVLMAQTLEKIFLOKVASMP 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	605	100.0	801	BRD2_HUMAN	P25440 homo sapien
2	542	89.6	726	BRD3_HUMAN	Q15059 homo sapien
3	500	82.6	1362	BRD4_HUMAN	O60885 homo sapien
4	498	82.3	2038	FSH_DROME	P13709 drosophila
5	202.5	33.5	727	YK82_SCHPO	Q9hgp4 schizosacch
6	187	30.9	832	PCAF_HUMAN	Q9h2811 homo sapien
7	184.5	30.5	2441	CBP_MOUSE	P45481 mus musculus
8	182	30.1	439	GCNE_YEAST	Q03330 saccharomyc
9	181.5	30.0	2414	P300_HUMAN	Q09472 homo sapien
10	180.5	29.8	2442	CBP_HUMAN	Q92793 homo sapien
11	176	29.1	686	BDP1_YEAST	P35817 saccharomyc
12	166.5	27.5	1484	CES2_HUMAN	Q9bxf3 homo sapien
13	162	26.8	2056	CBP1_CAEEL	P34545 caenorhabdi
14	160.5	26.5	1051	TF1A_MOUSE	O64127 mus musculus
15	159.5	26.4	1050	TF1A_HUMAN	Q15164 homo sapien
16	158	26.1	837	GCL2_HUMAN	Q92930 homo sapien
17	155	25.6	830	GCL2_MOUSE	Q9fjd2 mus musculus
18	149	24.6	1972	BA2B_HUMAN	Q9uif8 homo sapien
19	148	24.5	2130	BA2B_CHICK	Q9del3 gallus gall
20	139.5	23.1	1483	BA1B_HUMAN	Q9u190 homo sapien
21	136.5	22.6	1556	BA1A_HUMAN	Q9nrl2 h bromodoma
22	136	22.5	1872	T2D1_HUMAN	P21675 homo sapien
23	134.5	22.2	1479	BA1B_MOUSE	Q92277 mus musculus
24	133	22.0	1127	TF1G_HUMAN	Q9upn9 homo sapien
25	129	21.3	1850	BA2A_MOUSE	Q9lye5 mus musculus
26	129	21.3	1878	BA2A_HUMAN	Q9uif9 homo sapien
27	128.5	21.2	2269	WDR9_HUMAN	O9nsi6 homo sapien
28	127.5	21.1	2065	T2D1_DROME	P51123 drosophila
29	127	21.0	1638	BRM_DROME	P25439 drosophila
30	125	20.7	1214	BRF3_HUMAN	Q9uld4 homo sapien
31	123.5	20.4	627	BA1A_XENLA	O8uvr5 xenopus lae
32	121.5	20.1	1703	SNF2_YEAST	P22082 saccharomyc
33	114.5	18.9	1332	SPT7_YEAST	P35177 saccharomyc

```

34 109 18.0 625 1 RSC4_YEAST
35 108 17.9 1058 1 BRD1_HUMAN
36 106 17.5 542 1 BRD1_SCHPO
37 101 16.7 1214 1 BRF1_HUMAN
38 98.5 16.3 1186 1 PKCB_HUMAN
39 95 15.7 733 1 YC40_HUMAN
40 94 15.5 1359 1 STH1_YEAST
41 91.5 15.1 1586 1 SN22_HUMAN
42 89 14.7 1647 1 SN24_HUMAN
43 88 14.5 889 1 RSC2_YEAST
44 86 14.2 928 1 RSC1_YEAST
45 84 13.9 1242 1 TBP7_CAEEL

```

ALIGNMENTS

```

RESULT 1
BRD2_HUMAN
ID BRD2_HUMAN STANDARD; PRT; 801 AA.
AC P25440; O00699; O00700; Q15310; Q969U4;
DT 01-MAY-1992 (Rel. 22, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bromodomain-containing protein 2 (RING3 protein) (O27.1.1.1).
GN BRD2 OR RING3 OR KIAA9001.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RX MEDLINE=92329974; PubMed=1352711;
RA Beck S., Hanson I., Kelly A., Pappin D.J.C., Trowsdale J.;
RT "A homologue of the Drosophila female sterile homeotic (fsh) gene in
RL the class II region of the human MHC.";
RN [2]
RP DNA Seq. 2:203-210(1992).
RX REVISIONS TO N-TERMINUS.
RA MEDLINE=96376536; PubMed=8781126;
RA Thorpe K.L., Abdulla S., Kaufman J., Trowsdale J., Beck S.;
RT "Phylogeny and structure of the RING3 gene.";
RN [3]
RP Immunogenetics 44:391-396(1996).
RX SEQUENCE FROM N.A.
RA TISSUE=Bone marrow;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
RA Sato S., Nagase T., Seki T., Ishikawa K.-I., Tabata S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Thorpe K.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 2 bromodomains.
CC -!- SIMILARITY: Contains 1 ET domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X62083; CAA43996.1; -
CC EMBL; M80613; AAA68890.1; ALT_INIT.
CC EMBL; P42040; BAA07641.1; -
CC EMBL; Z84497; CAC69991.1; -
CC EMBL; Z96104; CAC69989.1; -
CC EMBL; X96670; CAA65450.1; -
CC HSSP; Q92831; 1B91.

```

```

CC      Genew; HGNC:1103; BRD2.
CC      MIM; 601540; -.
CC      GO; GO:0004574; F:protein serine/threonine kinase activity; TAS.
CC      GO; GO:0007283; P:spERMatogenesis; TAS.
CC      InterPro; IPR001487; Bromodomain.
CC      Pfam; PF00439; bromodomain; 2.
CC      PRINTS; PR00503; BROMODOMAIN.
CC      SMART; SM00297; BROMO; 2.
CC      PROSITE; PS00633; BROMODOMAIN 1; 2.
CC      PROSITE; PS50014; BROMODOMAIN 2; 2.
CC      KX Bromodomain; Repeat; Nuclear protein.
CC      FT DOMAIN 91 163 BROMODOMAIN 1.
CC      FT DOMAIN 364 436 BROMODOMAIN 2.
CC      FT DOMAIN 476 515 GLU/SER-RICH.
CC      FT DOMAIN 544 566 ARG/LYS-RICH (HIGHLY BASIC).
CC      FT DOMAIN 775 801 SER-RICH.
CC      FT DOMAIN 555 559 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC      FT DOMAIN 638 801 ET DOMAIN.
CC      FT DOMAIN 61 64 POLY-PRO.
CC      FT DOMAIN 492 506 POLY-GLU.
CC      FT DOMAIN 551 559 POLY-LYS.
CC      FT DOMAIN 634 638 POLY-GLU.
CC      FT DOMAIN 775 793 POLY-SER.
CC      FT CONFLICT 238 238 L -> F (IN REF. 4).
CC      SQ SEQUENCE 801 AA; 88060 MW; 9A075EBE13507D8E CRC64;

Query Match 100.0%; Score 605; DB 1; Length 801;
Best Local Similarity 100.0%; Pred. No. 5.4e-57; Indels 0; Gaps 0;
Matches 111; Conservative 0; Mismatches 0;

QY 1 KPGRTNQLYHLKVKYKALWKHGFAPFPQVDAVKLGLPDYHKILKQPMDMGTIKRRL 60
DB 71 KPGRTNQLYHLKVKYKALWKHGFAPFPQVDAVKLGLPDYHKILKQPMDMGTIKRRL 130
QY 61 ENNYWAASECMQDFTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 131 ENNYWAASECMQDFTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 181

RESULT 2
BRD3 HUMAN
ID BRD3 HUMAN STANDARD; PRT; 726 AA.
AC Q15059; Q92645;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Bromodomain-containing protein 3 (RING3-like protein).
GN BRD3 OR RING3L OR KIAA0043.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN TISSUE=Bone marrow;
RC SEQUENCE FROM N.A.
PX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kawabayashi Y., Ishikawa K.-I., Tabata S.;
RA "Prediction of the coding sequences of unidentified human genes. II.
RA The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RA analysis of cDNA clones from human cell line KG-1."
RA DNA Res. 1:223-229(1994).
[2]
RN SEQUENCE OF 363-726 FROM N.A.
PX MEDLINE=96038990; PubMed=9373153;
RA Thorpe K.L., Gorman P., Thomas C., Sheer D., Trowsdale J., Beck S.;
RA "Chromosomal localization, gene structure and transcription pattern of
RA the ORFX gene, a homologue of the MHC-linked RING3 gene."
RA Gene 200:177-183(1997).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: Contains 2 bromodomains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC ENBL; AR386649; AAL26987.1; -
CC HSP; Q92831; LB91.
CC DR Genew; HGNC:13575; BRD4.
CC DR InterPro; IPR001487; Bromodomain.
CC DR Pfam; PF00439; bromodomain; 2.
CC DR PRINTS; PR00503; BROMODOMAIN.
CC DR SMART; SM00297; BROMO; 2.
CC DR PROSITE; PS00633; BROMODOMAIN 1; 1.
CC DR PROSITE; PS00633; BROMODOMAIN 2; 2.
CC DR Bromodomain; Repeat; Nuclear protein.
CC FT DOMAIN 75 147 BROMODOMAIN 1.
CC FT DOMAIN 368 440 BROMODOMAIN 2.
CC FT DOMAIN 535 594 LYS-RICH.
CC FT DOMAIN 692 717 SER-RICH.
CC FT DOMAIN 703 714 POLY-SER.
CC FT DOMAIN 738 743 POLY-HIS.
CC FT DOMAIN 757 761 POLY-PRO.
CC FT DOMAIN 764 770 POLY-PRO.
CC FT DOMAIN 771 775 POLY-GLN.
CC FT DOMAIN 776 783 POLY-PRO.
CC FT DOMAIN 954 964 POLY-PRO.
CC FT DOMAIN 974 986 POLY-PRO.
CC FT DOMAIN 1011 1014 POLY-PRO.
CC FT DOMAIN 1028 1033 POLY-PRO.
CC FT DOMAIN 1283 1300 POLY-GLN.
CC FT DOMAIN 1301 1308 POLY-ALA.
CC FT DOMAIN 1335 1338 POLY-ARG.
CC FT CONFLICT 720 721 EM -> GP (IN REF. 2).
CC SQ SEQUENCE 1362 AA; 152219 MW; D52EFLCF9960907 CRC64;

Query Match 82.6%; Score 500; DB 1; Length 1362;
Best Local Similarity 80.2%; Pred. No. 1.6e-45;
Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVWKALWQHFAWPPQPDVAVKLGPDYHKLKIQPDMGTIKRL 60
DB 55 KPGRTNQLQYLRLVRLVLTWKHQFAWPPQPDVAVKLGPDYHKLKIQPDMGTIKRL 114
QY 61 ENNYWAASECMQDFNTMTNFCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 115 ENNYWAASECMQDFNTMTNFCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 165

RESULT 4
FSH DROME STANDARD; PRT; 2038 AA.
AC P13709; P13710;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Female sterile homeotic protein (Fragile-chorion membrane protein).
GS FS(1)H OR FSH.
OS Drosophila melanogaster (Fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC Ephyaroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89276730; PubMed=2567251;
RA Haynes S.R., Mozer B.A., Bhatia-Dey N., David I.B.;
RT "The Drosophila fish locus, a maternal effect homeotic gene, encodes
RT apparent membrane proteins";
RL Dev. Biol. 134:246-257 (1989).
CC -1- FUNCTION: Required maternally for proper expression of other
CC homeotic genes involved in pattern formation, such as UBX.
CC -1- SIMILARITY: HIGH, TO HUMAN RING3 PROTEIN.

CC -1- SIMILARITY: Contains 2 bromodomains.
CC -1- SIMILARITY: Contains 1 ET domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMEL; M23221; AAA28540.1; -
CC EMEL; M23222; AAA28541.1; ALT_TERM.
CC EMEL; M15762; AAA70424.1; -
CC EMEL; M15763; AAA70423.1; -
CC EMEL; M15764; AAA70422.1; -
CC PIR; A43742; A43742.
CC HSSP; Q92831; LB91.
CC FlyBase; FBgn0004656; fs(1)h.
CC InterPro; IPR001487; Bromodomain.
CC Pfam; PF00439; bromodomain; 2.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00297; BROMO; 2.
CC PROSITE; PS00633; BROMODOMAIN 1; 2.
CC PROSITE; PS00633; BROMODOMAIN 2; 2.
CC Developmental protein; Bromodomain; Transmembrane; Repeat.
CC FT DOMAIN 51 123 BROMODOMAIN 1.
CC FT DOMAIN 495 567 BROMODOMAIN 2.
CC FT DOMAIN 945 1106 ET DOMAIN.
CC FT TRANSMEM 330 350 POTENTIAL.
CC FT TRANSMEM 451 471 POTENTIAL.
CC FT TRANSMEM 750 770 POTENTIAL.
CC FT TRANSMEM 790 810 POTENTIAL.
CC FT TRANSMEM 816 830 POTENTIAL.
CC FT TRANSMEM 874 894 POTENTIAL.
CC FT TRANSMEM 1731 1751 POTENTIAL.
CC FT TRANSMEM 1939 1959 POTENTIAL.
CC FT VARIANT 909 909 G -> A.
CC FT VARIANT 1022 1022 H -> RXPYY.
CC SQ SEQUENCE 2038 AA; 205332 MW; 649E0706D50A0098 CRC64;

Query Match 82.3%; Score 498; DB 1; Length 2038;
Best Local Similarity 77.5%; Pred. No. 4e-45;
Matches 86; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVWKALWQHFAWPPQPDVAVKLGPDYHKLKIQPDMGTIKRL 60
DB 31 KPGRTNQLQYLKTVWKVWKHFWPPQPDVAVKLGPDYHKLKIQPDMGTIKRL 90
QY 61 ENNYWAASECMQDFNTMTNFCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
DB 91 ENNYWAASECMQDFNTMTNFCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 141

RESULT 5
YK82_SCHPO STANDARD; PRT; 727 AA.
AC Q9HGP4; P78808;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical bromodomain protein C631.02.
GN SPAC631.02.
OS Schizosaccharomyces pombe (Fission yeast).
CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,


```
DR GO; GO:0006338; P:chromatin modeling; NAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; IDA.
DR GO; GO:0006473; P:protein amino acid acetylation; TAS.
DR InterPro; IPR001487; GCMODOMAIN.
DR InterPro; IPR000182; GCSacetyltrans.
DR Pfam; PF00583; Acetyltransf; 1.
DR Pfam; PF00439; bromodomain; 1.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
KW Transcription regulation; Transferase; Nuclear protein;
KW Bromodomain; 3D-structure; Cell cycle.
FT DOMAIN 550 623 ACETYLTRANSFERASE.
FT MUTAGEN 740 810 BROMODOMAIN.
FT MUTAGEN 752 752 V-A: REDUCED ACETYL-LYSINE BINDING.
FT MUTAGEN 760 760 Y-A: REDUCED ACETYL-LYSINE BINDING.
FT MUTAGEN 802 802 Y-S: REDUCED ACETYL-LYSINE BINDING.
FT MUTAGEN 809 809 Y-S: COMPLETE LOSS OF ACETYL-LYSINE BINDING.
FT CONFLICT 716 718 KEK -> SHM (IN REF. 3).
FT CONFLICT 766 766 F -> S (IN REF. 3).
FT CONFLICT 805 805 A -> P (IN REF. 3).
SQ SEQUENCE 832 AA; 92960 MW; 72F516F8BD10CD0C CRC64;

Query Match 30.9%; Score 187; DB 1; Length 832;
Best Local Similarity 36.4%; Pred. No. 3e-12;
Matches 39; Conservative 18; Mismatches 48; Indels 2; Gaps 1;

QY 1 KPGVNTQLYKHKVMKALWKHQFAPRPQPDVAVKGLGPDYHKLIKQPMDMGTIKRL 60
DB 720 KEKRPDQLYKLSILQYKSHQSAWPFMEPVKRTIE--APGYEIVIRPMDLKTMSERL 777
QY 61 ENNYVWAASECMQDFNTFTNYCIYKPTDDIVLMAQTLEKIFLQKV 107
DB 778 KRYVYVKLFLMADLQRFVFNCKEYNAASEYKCANILEKPFPSKI 824

RESULT 7
CBP_MOUSE STANDARD; PRT; 2441 AA.
AC P45481;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CREB-binding protein (EC 2.3.1.48).
GN CREBBP OR CBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94019866; PubMed=8413673;
RA Chivria J.C., Kwok R.P.S., Lamb N., Hagiwara M., Montminy M.R.,
RA Goodman R.H.;
RT "Phosphorylated CREB binds specifically to the nuclear protein CBP.";
RL Nature 365:855-859(1993).
[2]
RP INTERACTION WITH NCOA3.
RX MEDLINE=97336097; PubMed=9192892;
RA Torchia J., Rose D.W., Inostroza J., Kamei Y., Westin S., Glass C.K.,
RA Rosenfeld M.G.;
RT "The transcriptional co-activator p/CIP binds CBP and mediates
RT nuclear-receptor function.";
RL Nature 387:677-684(1997).
CC -!- FUNCTION: Acetyltransferase enzyme. Acetylates histones, giving a
CC specific tag for transcriptional activation. Also acetylates non-
CC histone proteins, like NCOA3 coactivator. Mediates cAMP-gene
CC regulation by binding specifically to phosphorylated CREB protein.
CC CBP, as coactivator, augments the activity of phosphorylated CREB
CC to activate transcription of cAMP-responsive genes (By
```

Db 1099 LEALYRQDPESLPFRQVDPQLGIPDYFDIVKXPMDLSTIKRKLDTGQYQEPNQYDDV 1158

OY 76 NTMFNFCVYIKPTDIDVLVMAQTLKFLQKV 107

Db 1159 RLNFNNAWLNKRTSRVYKFCSCLAEPVEQEII 1190

RESULT 8

GCN5_YEAST

AC Q03330; STANDARD; PRT; 439 AA.

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Histone methyltransferase GCN5 (EC 2.3.1.48).

GN GCN5 OR ADA4 OR YOR252W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]_SEQUENCE FROM N.A.

RP MEDLINE=93011009; PubMed=1396595;

RX Georgakopoulos T., Thireos G.;

RA "Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription.";

RT EMBO J. 11:4145-4152(1992).

RL [2]

RP SEQUENCE FROM N.A.

RP MEDLINE=97313265; PubMed=9169869;

RX Tettelin H., Agostoni Carbone M.L., Albermann K., Albers K., Arroyo K., Backes U., Barreiros T., Bertani I., Bjourson A.J., Bruckner M., Bruschi C.V., Carignani G., Castagnoli L., Cerdan E., Clemente M.L., Coblenz A., Coglievina M., Coissac E., Defoor E., Del Bino S., Delius H., Delner D., de Wergifosse P., Dujon B., Durand P., Entian K.-D., Eraso P., Escribano V., Fabiani L., Fartmann B., Feroli P., Feuermann M., Frontali L., Garcia-Gonzalez M., Garcia-Saez M.I., Goffeau A., Guerreiro P., Hani J., Hansen M., Hebling U., Hernandez K., Heumann K., Hilger F., Hofmann B., Indge K.J., James C.M., Klima R., Koetter P., Kramer B., Kramer W., Laugin G., Leuther H., Louis E.J., Maillier E., Marconi A., Martegani E., Mason M.J., Mazzoni C., McReynolds A.D.K., Melchiorre P., Meves H.-W., Minenkova O., Mueller-Auer S., Naurrocki A., Netter P., Neu R., Nombela C., Oliver S.G., Panzeri L., Pazuoli S., Plevani P., Portetelle D., Porcillo F., Potier S., Purnelle B., Rieger M., Riles L., Rinaldi T., Robben J., Rodrigues-Pousada C., Rodriguez-Belmonte E., Rodriguez-Torres A.M., Rose M., Ruzzi M., Saliola M., Sanchez-Perez M., Schaefer B., Schaefer M., Schaefer M., Schmidheini T., Schreer A., Skala J., Souciet J.-L., Steensma H.Y., Talla E., Thierry A., Vandenbol M., Van der Aart Q.J.M., Van Dyck L., Vanoni M., Verhasselt P., Voet M., Volckaert G., Wambutt R., Watson M.D., Weber N., Wedler E., Wedler H., Wipfli P., Wolf K., Wright L.P., Zaccaria P., Zimmermann M., Zollner A., Kleine K.;

RT "The nucleotide sequence of Saccharomycetes cerevisiae chromosome VII.";

RL Nature 387:81-84(1997).

RN [3]

RP SEQUENCE OF 1-170 FROM N.A.

EC STRAIN=S288C / FY1679;

RX MEDLINE=9729234; PubMed=9133742;

RA Feroli F., Carignani G., Favanello A., Guerreiro P., Azevedo D., Rodrigues-Pousada C., Melchiorre P., Panzeri L., Agostoni Carbone M.L.;

RT "Analysis of a 17.9 kb region from Saccharomycetes cerevisiae chromosome VII reveals the presence of eight open reading frames, including BRF1 (TFIIIB70) and GCN5 genes.";

RL Yeast 13:373-377(1997).

RN [4]

RP SEQUENCE OF 170-439 FROM N.A.

EC STRAIN=S288C;

RX MEDLINE=9729233; PubMed=9133741;

RA Mazzoni C., Ruzzi M., Rinaldi T., Solinas F., Montebove F., Frontali L.;

RT "Sequence analysis of a 10.5 kb DNA fragment from the yeast chromosome VII reveals the presence of three new open reading frames and of a TRNATH gene.";

RL Yeast 13:369-372(1997).

RN [5]

RP ASSOCIATION WITH ADA2.

RX MEDLINE=95045371; PubMed=7957049;

RA Marcus G.A., Silverman N., Berger S.L., Horiuchi J., Guarente L.;

RT "Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors.";

RL EMBO J. 13:4807-4815(1994).

RN [6]

RP IDENTIFICATION IN A SAGA COMPLEX WITH SPT2; SPT7; SPT8; SPT20; HF11

ADA2; ADA3 AND TRAI.

RX MEDLINE=99102959; PubMed=9885573;

RA Grant P.A., Schielfitz D., Pray-Grant M.G., Yates J.R. III, Workman J.L.;

RT "The ATM-related cofactor Tra1 is a component of the purified SAGA complex.";

RL Mol. Cell 2:863-867(1998).

RN [7]

RP REVIEW.

RX MEDLINE=20400991; PubMed=10940244;

RA Dyda F., Klein D.C., Hickman A.B.;

RT "GCN5-related N-acetyltransferases: a structural overview.";

RL Annu. Rev. Biophys. Biomol. Struct. 29:81-103(2000).

RN [8]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 99-262.

RX MEDLINE=99362588; PubMed=10430873;

RA Trivelp R.C., Rojas J.R., Sterner D.E., Venkataramani R.N., Wang L., Zhou J., Allis C.D., Berger S.L., Marmorstein R.;

RT "Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:8931-8936(1999).

CC -!- FUNCTION: Acetylates Lys-14 of histone H3. Also acetylates Lys-8 and Lys-16 of histone H4 with a lower preference. Acetylation of histones gives a specific tag for epigenetic transcription activation. Operates in concert with certain DNA-binding transcriptional activators such as GCN4 or HAP2/3/4. Acts via the formation of large multiprotein complexes that modify the chromatin.

CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.

CC -!- SUBUNIT: Interacts with ADA2. Part of the ADA/GCN5 complex that consists of HFI1/ADA1, ADA2, ADA3, SPT20/ADA5 and GCN5. Component of the SAGA complex, at least composed of SPT2, SPT7, SPT8, SPT20/ADA5, HFI1, ADA2, ADA3/NGG1, TRA1 and GCN5.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- SIMILARITY: Contains 1 bromodomain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X68628; CAA48602.1; -.

CC EMBL; Z73037; CAA97281.1; -.

CC EMBL; X99228; CAA67614.1; -.

CC PIR; S28051; S28051.

CC PDB; 1YCH; 02-AUG-99.

CC PDB; 156I; 24-NOV-00.

CC GeneOnline; 141564; -.

CC TRANSFAC; T02145; -.

CC SGD; S0003484; GCN5.

CC GO; GO:0000124; C:SAGA complex; IDA.

CC InterPro; IPR01487; Bromodomain.

CC InterPro; IPR00182; GCN5acetyl_trans.

CC Pfam; PF00583; Acetyltransf; 1.

CC Pfam; PF00439; Bromodomain; 1.

CC PRINTS; PR00503; BROMODOMAIN.

CC SMART; SM00297; BROMO; 1.

DR PROSITE; PS00633; BROMODOMAIN 1; 1.
DR PROSITE; PS50014; BROMODOMAIN 2; 1.
KW Transcription regulation; Transferrin; Activator; Chromatin regulator;
KW Trans-acting factor; Nuclear protein; Bromodomain; 3D-structure.
FT ACT_SITE 173 173
FT DOMAIN 344 414
FT STRAND 100 105
FT STRAND 111 127
FT HELIX 129 130
FT TURN 133 141
FT TURN 143 144
FT STRAND 146 152
FT TURN 153 155
FT STRAND 156 166
FT STRAND 167 169
FT TURN 170 170
FT STRAND 171 179
FT TURN 181 182
FT TURN 186 187
FT HELIX 188 203
FT STRAND 208 213
FT HELIX 215 217
FT HELIX 218 223
FT TURN 224 225
FT STRAND 227 227
FT STRAND 234 237
FT STRAND 238 238
FT TURN 239 240
FT STRAND 241 241
FT TURN 244 245
FT STRAND 248 253
SQ SEQUENCE 439 AA; 51069 MW; 3200730DDC7EF70D CRC64;

Query Match 30.1%; Score 182; DB 1; Length 439;
Best Local Similarity 38.5%; Pred. No. 5.3e-12;
Matches 37; Conservative 16; Mismatches 41; Indels 2; Gaps 1;

QY 16 VMKALWKHQFANPPQVDVAVKGLPDYKIIQPMDMGTIKRLENNYWAASECMQDF 75
Db 339 ILTELQNHAAWPFQWV--KEVPDYDFIKPMDLSTMEIKLSNRYQKQMEDFIYDA 396
QY 76 NTNFTNCTYCNKPTDDIVLVAQTLKIFLOKVASMP 111
Db 397 RLVFNNCRWNGENTSYYKIANRLEKFFNNKVKEIP 432

RESULT 9
P300 HUMAN STANDARD; PRT; 2414 AA.
AC Q09472;
DI 15-JUN-1998 (Rel. 36, Created)
DI 15-JUL-1998 (Rel. 36, Last sequence update)
DI 10-OCT-2003 (Rel. 42, Last annotation update)
DE EIA-associated protein p300 (EC 2.3.1.48).
GN EP300 OR P300.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95011587; PubMed=7523245;
RA Eckner R., Eben M.E., Newsome D., Gerdes M., Decaprio J.A.,
RA Lawrence J.B., Livingston D.M.,
RT "Molecular cloning and functional analysis of the adenovirus EIA-
RT associated 300-KD protein (p300) reveals a protein with properties of
RT a transcriptional adaptor.";
RL Genes Dev. 8:869-884(1994).
RN [2]
RP ENZYMOLOGICAL ACTIVITY.
RX MEDLINE=97100994; PubMed=8945521;
RA Ogryzko V.V., Schiltz R.L., Russanova V., Howard B.H., Nakatani Y.,
RT "The transcriptional coactivators p300 and CBP are histone

RT acetyltransferases";
RL Cell 87:953-959(1996).
RN [3]
RP INTERACTION WITH PCAF.
RX MEDLINE=96300317; PubMed=8684459;
RA Yang X.-J., Ogryzko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.,
RT "A p300/CBP-associated factor that competes with the adenoviral
RT oncoprotein E1A.";
RL Nature 382:319-324(1996).
RN [4]
RP INTERACTION WITH HIF1A AND CREBBP.
RX MEDLINE=97075102; PubMed=8917528;
RA Arany Z., Huang L.E., Eckner R., Bhattacharya S., Jiang C.,
RA Goldberg M.A., Bunn H.F., Livingston D.M.,
RT "An essential role for p300/CBP in the cellular response to hypoxia.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12969-12973(1996).
RN [5]
RP INTERACTIONS WITH NR3C1.
RX MEDLINE=98250578; PubMed=9590696;
RA Fryer C.J., Archer T.K.,
RT "Chromatin remodelling by the glucocorticoid receptor requires the
RT BRG1 complex.";
RL Nature 393:88-91(1998).
RN [6]
RP INTERACTION WITH NCOA6.
RX MEDLINE=20283976; PubMed=10823961;
RA Ko L., Cardona G.R., Chin W.W.,
RT "Thyroid hormone receptor-binding protein, an LXXLL motif-containing
RT protein, functions as a general coactivator.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6212-6217(2000).
RN [7]
RP INTERACTION WITH DTX1.
RX MEDLINE=21576166; PubMed=11564735;
RA Yamamoto N., Yamamoto S.-I., Inagaki F., Kawauchi M., Fukamizu A.,
RA Kishi N., Matsuno K., Nakamura K., Weinmaster G., Okano H.,
RA Nakafuku M.,
RT "Role of Deltex-1 as a transcriptional regulator downstream of the
RT Notch receptor.";
RL J. Biol. Chem. 276:45031-45040(2001).
RN [8]
RP PHOSPHORYLATION, AND INTERACTION WITH TCF7L2 AND LEF1.
RX MEDLINE=22450614; PubMed=12446687;
RA Hecht A., Stemmler M.P.,
RT "Identification of a promoter-specific transcriptional activation
RT domain at the C-terminus of the Wnt-effector protein TCF4.";
RL J. Biol. Chem. 278:3776-3785(2003).
RN [9]
RP VARIANTS BREAST CANCER PRO-827 AND GLY-1013, VARIANT PANCREATIC CANCER
TYR-1650, AND VARIANT COLORECTAL CANCER GLN-2221.
RX MEDLINE=20164332; PubMed=10700188;
RA Gayther S.A., Batley S.J., Linger L., Bannister A., Thorpe K.,
RA Chin S.-F., Daigo Y., Russell P., Wilson A., Soutter H.M.,
RA Delhanty J.D.A., Ponder B.A.J., Kouzarides T., Caldas C.,
RT "Mutations truncating the EP300 acetylase in human cancers.";
RL Nat. Genet. 24:300-303(2000).
RN [10]
RP STRUCTURE BY NMR OF 302-418 IN COMPLEX WITH 786-826 OF HIF1A.
RX MEDLINE=21957254; PubMed=11953990;
RA Freedman S.J., Sun Z.-Y., Poy F., Kung A.L., Livingston D.M.,
RA Wagner G., Eck M.J.,
RT "Structural basis for recruitment of CBP/p300 by hypoxia-inducible
RT factor-1 alpha.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:5367-5372(2002).
CC -!- FUNCTION: Functions as histone acetyltransferase and regulates
CC transcription via chromatin remodeling. Acetylates all four core
CC histones in nucleosomes. Histone acetylation gives an epigenetic
CC tag for transcriptional activation. Binds to and may be involved
CC in the transforming capacity of the adenovirus E1A protein.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.
CC -!- SUBUNIT: Interacts with NR3C1, PCAF, NCOA6, LEF1 and TCF7L2. The
CC TAZ-type 1 domain interacts with HIF1A. Probably part of a complex
CC with HIF1A and CREBBP. Interacts with DTX1.
CC -!- SUBCELLULAR LOCATION: Nuclear.

GO; GO:0005634; C:nucleus; IDA.
GO; GO:0030528; F:transcription regulator activity; IMP.
GO; GO:0007151; P:sporation (sensu Saccharomycetes); IMP.
InterPro; IPR001487; Bromodomain.
Pfam; PF00439; bromodomain; 2.
PRINTS; PR00503; BROMODOMAIN.
SMART; SM00287; BROMO; 2.
PROSITE; PS00633; BROMODOMAIN_1; 2.
PROSITE; PS50014; BROMODOMAIN_2; 2.
KW Transcription regulation; Nuclear protein; Bromodomain; Repeat; Spoolation.
KW DOMAIN 165 237 BROMODOMAIN 1.
FT DOMAIN 332 404 BROMODOMAIN 2.
FT DOMAIN 522 586 ET DOMAIN.
FT CONFLICT 8 8 Q -> LC (IN REF. 1).
FT CONFLICT 93 94 GA -> R (IN REF. 2).
FT CONFLICT 94 94 A -> P (IN REF. 2).
FT CONFLICT 282 282 A -> P (IN REF. 1).
FT CONFLICT 385 385 D -> E (IN REF. 1).
FT CONFLICT 493 493 A -> R (IN REF. 4).
FT CONFLICT 493 493 A -> R (IN REF. 4).
SEQUENCE 586 AA; 76978 MW; 8CCD52F41F91D0DA CRC64;
Query Match 29.1%; Score 176; DB 1; Length 586;
Best Local Similarity 31.0%; Pred. No. 3.7e-11;
Matches 35; Conservative 25; Mismatches 51; Indels 2; Gaps 38
QY 1 KPGRTVNGLOGLHKVVKALKKH-QFAWPPRPQVDAVKGLGLPDYHKIIKQPMDMGTIKR 38
DB 310 KSKRLQQAQKFCQSVKELMAKKAASNYNPLEPVPDVSNNLPTFYDYVKEPMDLGTIAK 366
QY 59 RLENNYYWAASECQDMFTMTFCNYINKPDDIVLMAQTLEKIFLOKVASMP 111
DB 370 KLNDAQYQTMEDFFEDVRLVFKNCTYFNPDGITVNMGMGHRLEEVFNKWAADRP 422
RESULT 12
CES2 HUMAN
ID CES2 HUMAN STANDARD; PRT; 1484 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cat eye syndrome critical region protein 2.
GN CEK2 OR KIAA1740.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RP MEDLINE=21275466; PubMed=11381032;
RX Footz T.K., Brinkman-Mills P., Banting G.S., Maier S.A., Riazzi M.A.,
RA Brigland L.J., Hu S., Birren B., Minoshima S., Shimizu N., Pan H.,
RA Nguyen T., Fang Y., Fu Y., Ray L., Wu H., Shaull S., Phan S., Yao Z.,
RA Chen F., Huan A., Hu P., Wang Q., Loh P., Qi S., Roe B.A.,
RA McDermid H.E.;
RT "Analysis of the cat eye syndrome critical region in humans and the
RT region of conserved synteny in mice: a search for candidate genes at
RT or near the human chromosome 22 pericentromere.";
RL Genome Res. 11:1053-1070(2001).
RN [2]
RP SEQUENCE OF 346-1484 FROM N.A. (ISOFORM A).
RP TISSUE=Brain;
RC MEDLINE=21082932; PubMed=11214970;
RX Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355(2000).
RN [3]
RP SEQUENCE OF 168-1484 FROM N.A. (ISOFORM B), AND INTERACTION WITH
RP LRPRRC.
RP TISSUE=Liver;
RC

```

RX MEDLINE=21686162; PubMed=11827455;
RA Liu L., McKeenhan W.L.;
RT "Sequence analysis of LRPPRC and its SEC1 domain interaction partners
RT suggests roles in cytoskeletal organization, vesicular trafficking,
RT nucleocytoplasmic shuttling, and chromosome activity.";
RL Genomics 79:124-136(2002).
CC -!- FUNCTION: May be involved through its interaction with LRPPRC in
CC the integration of cytoskeletal network with vesicular
CC trafficking, nucleocytoplasmic shuttling, transcription, chromosome
CC remodeling and cytokinesis.
CC -!- SUBUNIT: Interacts with LRPPRC.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=Q9BXF3-1; Sequence=Displayed;
CC Name=B; Synonyms=CEC82B;
CC IsoId=Q9BXF3-2; Sequence=VSP_000571, VSP_000572, VSP_000573;
CC TISSUE SPECIFICITY: Highly expressed in skeletal muscle, thymus,
CC placenta and lung. Expressed at lower level in brain, heart,
CC colon, spleen, kidney.
CC -!- MISCELLANEOUS: Candidate gene for the Cat Eye Syndrome (CES), a
CC developmental disorder associated with the duplication of a 2 Mb
CC region of 22q11.2. Duplication usually takes in the form of a
CC supernumerary bisatellited isodicentric chromosome, resulting in
CC four copies of the region (represents an inv dup(22)(q11)). CES is
CC characterized clinically by the combination of coloboma of the
CC iris and anal atresia with fistula, downslanting palpebral
CC fissures, preauricular tags and/or pits, frequent occurrence of
CC heart and renal malformations, and normal or near-normal mental
CC development.
CC -!- SIMILARITY: Contains 1 bromodomain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF336133; AAK15343.1; -.
CC EMBL; AB051527; BAB21831.1; -.
CC EMBL; AF411609; AAL07393.1; -.
CC HSSP; Q92831; 1B91.
CC Genew; HGNC:1840; CEC82.
CC MIM; 607576; -.
CC GO; GO:0000910; P:cytokinesis; NAS.
CC GO; GO:0007010; P:cytoskeleton organization and biogenesis; NAS.
CC GO; GO:0016192; P:vesicle-mediated transport; NAS.
CC InterPro; IPR001487; Bromodomain.
CC Pfam; PF00439; bromodomain; 1.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00297; BROMO; 1.
CC PROSITE; PS00633; BROMODOMAIN 1; 1.
CC PROSITE; PS0014; BROMODOMAIN 2; 1.
CC Bromodomain; Alternative splicing.
KW DOMAIN 451 521 BROMODOMAIN.
FT DOMAIN 333 337 POLY-GLU.
FT DOMAIN 611 614 POLY-SER.
FT DOMAIN 1250 1253 POLY-PRO.
FT VARSPPLIC 291 318 Missing (in isoform B).
FT FTId=VSP_000571.
FT FTId=VSP_000572.
FT VARSPPLIC 519 526 EXTMSDN -> GKGRSLC (in isoform B).
FT FTId=VSP_000572.
FT VARSPPLIC 527 1484 Missing (in isoform B).
FT FTId=VSP_000573.
FT CONFLICT 370 389 MISSING (IN REF. 2).
FT CONFLICT 1029 1029 C -> S (IN REF. 2).
FT CONFLICT 1045 1045 R -> W (IN REF. 2).
SQ SEQUENCE 1484 AA; 164214 MW; 049AA844E51AF63F CRC64;
Query Match 27.5%; Score 166.5; DB 1; Length 1484;
Best Local Similarity 37.2%; Pred. No. 8.6e-10;

```

```

Matches 35; Conservative 21; Mismatches 31; Indels 7; Gaps 3;
QY 12 LHKV--VNKALWKEQFAPFPQPDVAVKLGDPYHKIKQPMGTTRRLNNYYWAAS 69
DB 443 MYKVLDDVVKV---HKDSWPFLEPVD--ESVAPNYQIKAPMDISSMEKKGLGYCTKE 497
QY 70 ECMODFTMTFNCYINYPKPTDDIVLMAQTLEKIF 103
DB 498 EFVNDMTKTFNCRKNGESSEYTKNSDNLRCF 531
RESULT 13
CPBL CAEEL STANDARD; PRT; 2056 AA.
ID CPBL CAEEL
AC P34545;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein cbp-1.
GN CBP-1 OR R10E11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden R.,
RA Stuston J., Thierley-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Worldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Duthin R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=B;
CC IsoId=P34545-1; Sequence=Displayed;
CC Name=A;
CC IsoId=P34545-2; Sequence=VSP_000557;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Contains 1 bromodomain.
CC -!- SIMILARITY: Contains 1 Z2-type zinc finger.
CC -!- SIMILARITY: Contains 2 TAZ-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z29095; CAAB2353.2; -.
CC EMBL; Z29095; CAD18875.1; -.
CC PIR; G88564; G88564.
CC WormPep; R10E11.1a; CE28069.
CC WormPep; R10E11.1b; CE21117.
CC InterPro; IPR001487; Bromodomain.
CC InterPro; IPR003101; KIX.
CC InterPro; IPR000197; TAZ_finger.

```

DR InterPro; IPR001965; Znf PHD.
 DR InterPro; IPR00433; Znf_ZZ.
 DR Pfam; PF00439; bromodomain; 1.
 DR Pfam; PF02172; KIX; 1.
 DR Pfam; PF02135; zf-TAZ; 2.
 DR Pfam; PF00569; ZF; 1.
 DR Pfam; PF00503; BROMODOMAIN.
 DR SMART; SM00297; BROMO; 1.
 DR SMART; SM00551; Znf TAZ; 2.
 DR SMART; SM00291; Znf_ZZ; 1.
 DR PROSITE; PS00633; BROMODOMAIN_1; 1.
 DR PROSITE; PS0014; BROMODOMAIN_2; 1.
 DR PROSITE; PS01359; ZF PHD; 1.
 DR PROSITE; PS0134; ZF TAZ; 2.
 DR PROSITE; PS01357; ZF_ZZ; 1.
 DR PROSITE; PS0135; ZF_ZZ_2; 1.
 KW Bromodomain; Metal-binding; Zinc; Zinc-finger; Alternative splicing;
 KW Repeat.
 FT ZN-FING 399 505 TAZ-TYPE 1.
 FT DOMAIN 881 953 BROMODOMAIN.
 FT ZN-FING 1493 1534 ZZ-TYPE.
 FT ZN-FING 1550 1631 TAZ-TYPE 2.
 FT DOMAIN 1687 2008 GLY/GLN-RICH.
 FT VARSPLIC 467 478 SDTCTTKKGSV -> F (in isoform a).
 FT SEQUENCE 2056 AA; 227179 MW; 949FF4608C634F01 CRC64;
 Query Match 26.8%; Score 162; DB 1; Length 2056;
 Best Local Similarity 35.9%; Pred. No. 3.7e-09;
 Matches 33; Conservative 19; Mismatches 40; Indels 0; Gaps 0;
 QY 16 VMKALWKHOFAPRPQVDVAVKGLPYHKIKQPMGDKIKRLENNYVWASECQDF 75
 DB 876 VWEKLDKSEDAAPRPVVDKLLNIPDHEIKRPMLETVHKLKYAGQYQNGQFCDDI 935
 QY 76 NTMTFCYIYNKPTDDIVLMAQTLEKIFLQKV 107
 DB 936 WLMLDNLWLNKNSKVYKGLKLESEFVSEM 967
 RESULT 14
 TFIIA_MOUSE STANDARD; PRT; 1051 AA.
 ID AC Q64127; Q64126;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Transcription intermediary factor 1-alpha (TFIIA-alpha) (Tripartite
 DE motif protein 24).
 GN TFII OR TRIM24 OR TFIIA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Carcinoma;
 RX MEDLINE=95262642; PubMed=7744009;
 RA le Douarin B., Zechel C., Garnier J.-M., Lutz Y., Tora L., Pierrat B.,
 RA Heery D., Gronemeyer H., Chambon P., Lessou R.;
 RT "The N-terminal part of TFIIA, a putative mediator of the ligand-
 RT dependent activation function (AF-2) of nuclear receptors, is fused to
 RT B-raf in the oncogenic protein TIF-2.";
 RL EMBL J. 14:2020-2033(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherf A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP INTERACTION WITH CBX1 AND CBX3.
 RX MEDLINE=97133299; PubMed=8978696;
 RA le Douarin B., Nielsen A.L., Garnier J.-M., Ichinose H.,
 RA Jeanmougin F., Lessou R., Chambon P.;
 RT "A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic
 RT control of transcription by nuclear receptors.";
 RL EMBL J. 15:6701-6715(1996).
 CC -!- FUNCTION: Interacts selectively in vitro with the AF2-activating
 CC domain of the estrogen receptors. Association with DNA-bound
 CC estrogen receptors requires the presence of estradiol (By
 CC similarity).
 CC -!- SUBUNIT: Interacts with NR3C2 (By similarity). Interacts with CBX1
 CC and CBX3.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q64127-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=Q64127-2; Sequence=VSP_005773;
 CC -!- TISSUE SPECIFICITY: Detected in all adult tissues, with the
 CC highest expression level in testis.
 CC -!- DISPAR: Participates in a chromosomal translocation that produces
 CC a TIF1A-BRAF (T19) oncogene originally isolated from a furfural-
 CC induced hepatoma.
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -!- SIMILARITY: Contains 2 B box-type zinc fingers.
 CC -!- SIMILARITY: Contains 1 bromodomain.
 CC -!- SIMILARITY: Contains 1 PHD-type zinc finger.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S78221; AAB34290.1; -.
 CC EMBL; S78219; AAB34289.1; -.
 CC EMBL; BC056959; AAB56959.1; -.
 CC PIR; S55259; S55259.
 CC HSSP; P29590; 1BOR.
 CC TRANSFAC; T02143; -.
 CC TRANSFAC; T02215; -.
 CC MGD; MGI:109275; Trim24.
 CC GO; GO:0005737; Cytoplasm; IDA.
 CC InterPro; IPR003849; Bbox_C.
 CC InterPro; IPR001487; Bromodomain.
 CC InterPro; IPR000315; Znf Bbox.
 CC InterPro; IPR001965; Znf PHD.
 CC InterPro; IPR001841; Znf_ring.
 CC Pfam; PF00439; bromodomain; 1.
 CC Pfam; PF00628; PHD; 1.
 CC Pfam; PF00543; zf-B_Box; 2.
 CC Pfam; PF00097; zf-C3HC4; 1.
 CC PRINTS; PR01406; BBOXZNFINGER.

```

DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00502; BBC; 1.
DR SMART; SM00336; BBOX; 2.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00249; PHD; 2.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
DR PROSITE; PS00119; ZF_BBOX; 2.
DR PROSITE; PS01359; ZF_PHD_1; 1.
DR PROSITE; PS00016; ZF_PHD_2; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00099; ZF_RING_2; 1.
DR TRANSCRIPTION REGULATION; Repressor; DNA-binding; Bromodomain;
KW Zinc-finger; Alternative splicing; Nuclear protein; Coiled coil;
KW Repeat; Proto-oncogene; Chromosomal translocation.
FT DOMAIN 8 15 POLY-ALA.
FT DOMAIN 19 22 POLY-ALA.
FT ZN_FING 52 77 RING-TYPE.
FT ZN_FING 158 211 B_BOX-TYPE 1.
FT ZN_FING 218 259 B_BOX-TYPE 2.
FT DOMAIN 289 359 COILED COIL (POTENTIAL).
FT DOMAIN 344 347 POLY-GLN.
FT DOMAIN 583 587 POLY-SER.
FT DOMAIN 755 780 NUCLEAR RECEPTOR BINDING SITE (NRBS).
FT ZN_FING 827 874 PHD-TYPE.
FT DOMAIN 892 908 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 933 988 BROMODOMAIN.
FT SITE 332 333 BREAKPOINT FOR TRANSLOCATION TO FORM
TIF1A-BRAF ONCOGENE.
FT VARSPLIC 477 510 Missing (in isoform Short).
SQ SEQUENCE 1051 AA; 116656 MW; 610584f3c6885972 CRC64;
Query Match 26.5%; Score 160.5; DB 1; Length 1051;
Best Local Similarity 38.4%; Pred. No. 2.6e-09;
Matches 33; Conservative 17; Mismatches 31; Indels 5; Gaps 2;
QY 20 LMKHQPWFPPVDVAVKGLGPDYHKIKQPMDMGTIKERLENNY--YWAASECQMDFNT 77
Db 916 LYCHENSFLAFQDP--VPLTVDPDYKIKNPMDLSIKKLQEDICMYTKPDEFDVDFRL 972
QY 78 MFTNCYIYNKPTDDIVLMAQTLEKIF 103
Db 973 IFQNCAEFNEPDPSEVANAGIKLESYF 998
RESULT 15
TIF1A_HUMAN STANDARD; PRT; 1050 AA.
AC O15164; O93854;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription intermediary factor 1-alpha (TIF1-alpha) (Tripartite
DE motif protein 24)
GN TIF1 OR TRIM24 OR TIF1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Breast cancer;
RX MEDLINE=97277352; PubMed=9115274;
RA Therot S., Henriquet C., Rochefort H., Cavailles V.;
RT "Differential interaction of nuclear receptors with the putative human
RT transcriptional coactivator hTIF1";
RL J. Biol. Chem. 272:12062-12068(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=99144725; PubMed=10022127;
RA Venturini L., You J., Stadler M., Gallien R., Lallemand V.,

```

```

RA Koken M.H.M., Mattei M.-G., Ganser A., Chambon P., Losson R.,
RA De The H.;
RT "TIF1gamma, a novel member of the transcriptional intermediary factor
RT 1 family.";
RL Oncogene 18:1209-1217(1999).
RN [3]
RP SEQUENCE OF 477-510 (ISOFORM LONG).
RC TISSUE=Breast cancer;
RA Cavailles V.;
RL Submitted (JAN-1999) to Swiss-Prot.
RN [4]
RP INTERACTION WITH NR3C2.
RX MEDLINE=2038588; PubMed=10935545;
RA Hellal-Levy C., Pagart J., Souque A., Wurtz J.-M., Moras D.,
RA Rafestin-Oblin M.-E.;
RT "Crucial role of the H11-H12 loop in stabilizing the active
RT conformation of the human mineralocorticoid receptor.";
RL Mol. Endocrinol. 14:1210-1221(2000).
CC -!- FUNCTION: Interacts selectively in vitro with the AF2-activating
CC domain of the estrogen receptors. Association with DNA-bound
CC estrogen receptors requires the presence of estradiol.
CC -!- SUBUNIT: Interacts with CBX1 and CBX3 (By similarity). Interacts
CC with NR3C2.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=O15164-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=O15164-2; Sequence=VSP_005772;
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 2 B box-type zinc fingers.
CC -!- SIMILARITY: Contains 1 bromodomain.
CC -!- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF009353; AAB63585.1; -.
DR EMBL; AF119042; AAD17258.1; -.
DR HSSP; P29590; 1BCR.
DR TRANSPAC; T04945; -.
DR Genew; HGNC:11812; TIF1.
DR MIM; 603406; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0005102; F:receptor binding; TAS.
DR GO; GO:0003713; F:transcription co-activator activity; TAS.
DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.
DR InterPro; IPR003649; BBox_C.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR000315; Znf_Bbox.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00439; bromodomain; 1.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF00643; zf-B_box; 2.
DR Pfam; PF00097; zf-C3HC4; 1.
DR PRINTS; PR01406; BBOXZNFINGER.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00502; BBC; 1.
DR SMART; SM00336; BBOX; 2.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
DR PROSITE; PS01359; ZF_BBOX; 2.
DR PROSITE; PS01359; ZF_PHD_1; 1.

```

DR PROSITE; PS0016; ZF_PHD_2; 1.
DR PROSITE; PS0018; ZF_RING_1; 1.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Transcription regulation; Repressor; DNA-binding; Bromodomain;
KW Zinc-finger; Alternative splicing; Nuclear protein; Coiled coil;
KW Repeat.
FT DOMAIN 9 15 POLY-ALA.
FT ZN_FING 56 82 RING-TYPE.
FT ZN_FING 158 211 B BOX-TYPE 1.
FT ZN_FING 218 259 B BOX-TYPE 2.
FT DOMAIN 289 359 COILED COIL (POTENTIAL).
FT DOMAIN 344 347 POLY-GLN.
FT DOMAIN 754 779 NUCLEAR RECEPTOR BINDING SITE (NRBS).
FT ZN_FING 826 873 PHD-TYPE.
FT DOMAIN 891 907 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 932 987 BROMODOMAIN.
FT VARSPLIC 477 510 Missing (in isoform Short).
FT CONFLICT 14 20 AASAAAS -> RLGCAP (IN REF. 1).
FT CONFLICT 24 28 SAAPS -> RGG (IN REF. 1).
FT CONFLICT 109 114 GSPVSG -> ARRSA (IN REF. 1).
FT CONFLICT 350 350 A -> T (IN REF. 1).
FT CONFLICT 600 600 D -> N (IN REF. 1).
FT CONFLICT 608 608 M -> I (IN REF. 1).
FT CONFLICT 967 967 A -> R (IN REF. 1).
SQ SEQUENCE 1050 AA; 116831 MW; D341E8022AAC67E CRC64;
Query Match 26.4%; Score 159.5; DB 1; Length 1050;
Best Local Similarity 38.4%; Pred.No. 3.4e-09;
Matches 33; Conservative 17; Mismatches 31; Indels 5; Gaps 2;
QY 20 LWKQFAPFPQPDVAVKGLPDYHKIKOPMDMGTKRLENNY--YWAASECMQDFNT 77
DB 915 LYCHEMSLAFODP--VPLTVPDYIKIKNPMDLUSTIKKRQEDYSWKPDVFADRL 971
QY 78 MFTNCYIYNKPTDDIVLMAQTLERIF 103
DB 972 IFQCAEFNEPDSEVANAGIKLENYF 997

Search completed: July 9, 2004, 13:06:42
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:03:30 ; Search time 39 Seconds
(without alignments)
898.013 Million cell updates/sec

Title: US-09-784-553C-19
Perfect score: 605
Sequence: 1 KGRVTNQYLHKVVMKAL.....IVLMAQTLEKFLQKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPRTREMBL25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605	100.0	503	11	Q99PC5 mus musculus
2	605	100.0	798	11	O8411 mus musculus
3	605	100.0	798	11	O54795 mus musculus
4	594	98.2	729	13	Q90971 gallus galli
5	561	92.7	539	13	Q72VH4 xenopus lae
6	542	89.6	505	11	O8C665 mus musculus
7	542	89.6	511	11	O8CAX7 mus musculus
8	542	89.6	556	4	Q8NSM3 homo sapien
9	542	89.6	726	11	O8KZF0 mus musculus
10	542	89.6	726	11	Q9J125 mus musculus
11	530	87.6	499	13	Q7SXM5 brachydanio
12	524	86.6	515	13	Q7SXL8 brachydanio
13	521	86.1	701	13	O8OPT7 fugu rubrip
14	517	85.5	558	13	Q7SXP6 brachydanio
15	500	82.6	723	11	O8VHF7 mus musculus
16	500	82.6	731	4	O60433 homo sapien

17	500	82.6	1400	11	Q9ESU6 mus musculus
18	500	82.6	1400	11	O8VHF8 mus musculus
19	500	82.6	1846	4	O86Y88 homo sapien
20	499	82.5	814	13	O8UUM2 deum2 oryzias lat
21	498	82.3	1110	5	O8IRN6 Q9313 drosophila
22	498	82.3	2038	5	O9W3L3 Q9313 drosophila
23	494	81.7	806	13	O8AWX9 Q9awx9 brachydanio
24	491	81.2	732	13	Q9GU61 Q9gu61 myxine glut
25	484	80.0	585	13	Q7ZVG9 Q7zvg9 brachydanio
26	482	79.7	664	5	O8T775 Q8t775 branchiost
27	459	75.9	463	4	O8IWI6 Q8iwi6 homo sapien
28	459	75.9	947	4	O14789 O14789 homo sapien
29	459	75.9	960	4	O7Z4A6 O7z4a6 homo sapien
30	449	74.2	956	11	Q91Y44 Q91y44 mus musculus
31	304	50.2	249	13	Q91686 Q91686 xenopus lae
32	296	48.9	1209	5	O20947 O20947 caenorhabdi
33	281.5	46.5	247	5	O8T3Z6 O8t3z6 drosophila
34	281.5	46.5	513	5	Q9VC56 Q9vc56 drosophila
35	280	46.3	765	5	O8ES79 Q8es79 caenorhabdi
36	280	46.3	851	5	O9SY80 Q9sy80 caenorhabdi
37	279	46.1	1087	5	O20948 O20948 caenorhabdi
38	268	44.3	109	13	O57335 O57335 brachydanio
39	237	39.2	674	5	Q9V8S1 Q9v8s1 drosophila
40	237	39.2	679	5	O8T3Z8 O8t3z8 drosophila
41	221.5	36.6	688	10	Q93YS6 Q93ys6 arabidopsis
42	221.5	36.6	703	10	Q9LYA2 Q9ly42 arabidopsis
43	215	35.5	578	3	Q9Y7N0 Q9y7n0 schizosacch
44	208	34.4	818	10	Q9LK27 Q9lk27 arabidopsis
45	204	33.7	386	10	O84XV2 O84xv2 arabidopsis

ALIGNMENTS

RESULT 1

Q99PC5 PRELIMINARY; PRT; 503 AA.

AC Q99PC5; 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Ring 3 (Fragment).

GN BRD2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL;

RA	Korf I.;
RT	"Complete sequence of UL26B06.";
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF318183; AAK07919.1; -.
DR	HSSP; Q92831; 1B91.
DR	MGI; 99495; Brd2.
DR	InterPro; IPR001487; Bromodomain.
DR	Pfam; PF00439; bromodomain; 2.
DR	PRINTS; PRO0503; BROMODOMAIN.
DR	SMART; SM00297; BROMO; 2.
DR	PROSITE; PS00633; BROMODOMAIN_1; 2.
DR	PROSITE; PS0014; BROMODOMAIN_2; 2.
FT	NON_TER 503 503
SQ	SEQUENCE 503 AA; 842A566963981F28 CRC64;

Query Match 100.0%; Score 605; DB 11; Length 503;
Best Local Similarity 100.0%; Fred. No. 5e-59;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGRVTNQYLHKVVMKALWKHQFAPFPQVDVAVKLGFLPDYHKIIKQPMDMGTIKRL 60
DB 70 KGRVTNQYLHKVVMKALWKHQFAPFPQVDVAVKLGFLPDYHKIIKQPMDMGTIKRL 129

QY 61 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
 DB 130 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 180

RESULT 2
 ID C88411 PRELIMINARY; PRT; 798 AA.
 AC C88411.
 DT 01-NOV-1998 (TRENBLrel. 08, Created)
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Female sterile homeotic-related protein Frg-1.
 GN BRD2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rhee K., Brunori M., Besset V., Wolgemuth D.J.;
 RT "Expression and potential role of Frg-1, a putative murine
 bromodomain-containing homologue of the Drosophila gene female sterile
 RT homeotic.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF045462; AAC24810.1; -.
 DR HSSP; Q92831; 1B91.
 DR MGD; MGI:99495; Brd2.
 DR InterPro; IPR001487; Bromodomain.
 DR PRINTS; PR00503; BROMODOMAIN.
 DR SMART; SM00297; BROMO; 2.
 DR PROSITE; PS00633; BROMODOMAIN_1; 2.
 DR PROSITE; PS00014; BROMODOMAIN_2; 2.
 SQ SEQUENCE 798 AA; 88063 MW; A9942517CF15B7A1 CRC64;

Query Match 100.0%; Score 605; DB 11; Length 798;
 Best Local Similarity 100.0%; Pred. No. 8.4e-59;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRVNQLQYLHKVVKALWKHQFAPFPQPDVAVKGLPDYHKIHKQPMDMGTIKRRL 60
 DB 70 KPGRVNQLQYLHKVVKALWKHQFAPFPQPDVAVKGLPDYHKIHKQPMDMGTIKRRL 129

QY 61 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
 DB 130 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 180

RESULT 3
 ID C54795 PRELIMINARY; PRT; 798 AA.
 AC C54795;
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE RING3 protein.
 GN BRD2 OR PSRG1 OR RING3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thorpe K.L., Beck S.;
 RL Immunogenetics 0:0-0(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SVJ;
 RA Taniguchi Y., Matsuzaka Y., Fujimoto H., Miyado K., Kohda A.,
 RA Okumura K., Kimura M., Inoko H.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE OF 47-549 FROM N.A.
 RC STRAIN=INBRED CD-1; TISSUE=Testis;
 RA Taniguchi Y., Matsuzaka Y., Fujimoto H., Miyado K., Kohda A.,
 RA Okumura K., Kimura M., Inoko H.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SVJ;
 RA Rowen L., Qin S., Madan A., Loretz C., James R., Dors M., Max L.,
 RA Hall J., Lasky S., Hood L.;
 RT "Sequence of the mouse major histocompatibility locus class II
 RT region.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL009226; CAA15818.1; -.
 DR EMBL; AL009226; CAA15819.1; -.
 DR EMBL; D89801; BAA25416.1; -.
 DR EMBL; AB010248; BAA24379.1; -.
 DR EMBL; AB010247; BAA24378.1; -.
 DR EMBL; AB010246; BAA24377.1; -.
 DR EMBL; AF100956; AAC69907.1; -.
 DR HSSP; Q92831; 1B91.
 DR MGD; MGI:99495; Brd2.
 DR InterPro; IPR001487; Bromodomain.
 DR Pfam; PF00439; Bromodomain; 2.
 DR PRINTS; PR00503; BROMODOMAIN.
 DR SMART; SM00297; BROMO; 2.
 DR PROSITE; PS00633; BROMODOMAIN_1; 2.
 DR PROSITE; PS00014; BROMODOMAIN_2; 2.
 SQ SEQUENCE 798 AA; 88066 MW; 08DD57FBF1385E96 CRC64;

Query Match 100.0%; Score 605; DB 11; Length 798;
 Best Local Similarity 100.0%; Pred. No. 8.4e-59;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRVNQLQYLHKVVKALWKHQFAPFPQPDVAVKGLPDYHKIHKQPMDMGTIKRRL 60
 DB 70 KPGRVNQLQYLHKVVKALWKHQFAPFPQPDVAVKGLPDYHKIHKQPMDMGTIKRRL 129

QY 61 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
 DB 130 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 180

RESULT 4
 ID Q90971 PRELIMINARY; PRT; 729 AA.
 AC Q90971;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Kinase.
 GN RING3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B12;
 RX MEDLINE=96376536; PubMed=8781126;
 RA Thorpe K.L., Abdulla S., Kaufman J., Trowsdale J., Beck S.;
 RT "Phylogeny and structure of the RING3 gene.";
 RL Immunogenetics 44:391-396(1996).
 DR EMBL; X96669; CAA65449.1; -.
 DR HSSP; Q92831; 1B91.
 DR InterPro; IPR001487; Bromodomain.
 DR Pfam; PF00439; Bromodomain; 2.
 DR PRINTS; PR00503; BROMODOMAIN.
 DR SMART; SM00297; BROMO; 2.
 DR PROSITE; PS00633; BROMODOMAIN_1; 2.
 DR PROSITE; PS00014; BROMODOMAIN_2; 2.
 SQ SEQUENCE 729 AA; 80389 MW; 06E1B92C804DF7B8 CRC64;

Query Match 98.2%; Score 594; DB 13; Length 729;
Best Local Similarity 97.3%; Pred. No. 1.3e-57;
Matches 108; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKKIKQPMDMGTIKRRL 60
DB 24 KPGRTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKKIKQPMDMGTIKRRL 83
QY 61 ENNYWAASECQDFTMTFNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 84 ENNYWGAECQDFTMTFNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 134

RESULT 5
Q7ZYH4 PRELIMINARY; PRT; 539 AA.
AC Q7ZYH4;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to bromodomain-containing 2 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC043784; AAH43784.1; -
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN 1; 1.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
FT NON_TER 539 539
SQ SEQUENCE 539 AA; 60130 MW; 53841988925415F4 CRC64;

Query Match 92.7%; Score 561; DB 13; Length 539;
Best Local Similarity 89.2%; Pred. No. 4.4e-54;
Matches 99; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKKIKQPMDMGTIKRRL 60
DB 71 KPGRTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKKIKQPMDMGTIKRRL 130
QY 61 ENNYWAASECQDFTMTFNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 131 ENNYWGAECQDFTMTFNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 181

RESULT 6
Q8C665 PRELIMINARY; PRT; 505 AA.
AC Q8C665;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Bromodomain-containing 3 (Fragment).
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
DR EMBL; AK076472; BAC36359.1; -
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN 1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
FT NON_TER 505 505
SQ SEQUENCE 505 AA; 55211 MW; 35C76A91AAEC6AB4 CRC64;

Query Match 89.6%; Score 542; DB 11; Length 505;
Best Local Similarity 87.4%; Pred. No. 5.4e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKKIKQPMDMGTIKRRL 60
DB 30 KPGRTNQLQYMNQVVKTLWKHQFAMFPFQPDVAVKLGDPYHKKIKQPMDMGTIKRRL 89
QY 61 ENNYWAASECQDFTMTFNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 90 ENNYWGAECQDFTMTFNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 140

RESULT 7
Q8CAX7 PRELIMINARY; PRT; 511 AA.
AC Q8CAX7;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Bromodomain-containing 3 (Fragment).
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
DR EMBL; AK037435; BAC29806.1; -
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN 1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
FT NON_TER 511 511
SQ SEQUENCE 511 AA; 55921 MW; E189A3484AEB2679 CRC64;

Query Match 89.6%; Score 542; DB 11; Length 511;
Best Local Similarity 87.4%; Pred. No. 5.5e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKKIKQPMDMGTIKRRL 60
DB 30 KPGRTNQLQYMNQVVKTLWKHQFAMFPFQPDVAVKLGDPYHKKIKQPMDMGTIKRRL 89
QY 61 ENNYWAASECQDFTMTFNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 90 ENNYWGAECQDFTMTFNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 140

```

RESULT 8
Q8NSM3 PRELIMINARY; PRT; 556 AA.
AC Q8NSM3;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to bromodomain containing 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032124; AAH32124.1; -.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; Bromodomain; 2.
DR PRINTS; PR00503; Bromodomain.
DR SMART; SM00297; BRMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS0014; BROMODOMAIN_2; 2.
SQ SEQUENCE 556 AA; 60942 MW; 8352F5DF1801A793 CRC64;

Query Match 89.6%; Score 542; DB 4; Length 556;
Best Local Similarity 87.4%; Pred. No. 6.1e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHVKKVVKALWKHGFQFAPFPQVDAVKLGLPDYHKIKQPMDMGTIKRL 60
DB 31 KPGRTNQLQYMNQVNVKTLWKHGFQFAPFPQVDAIKLNPDYHKIKQPMDMGTIKRL 90
QY 61 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 111
DB 91 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQALEKIFLQKVAQMP 141

RESULT 9
Q8K2F0 PRELIMINARY; PRT; 726 AA.
AC Q8K2F0;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Bromodomain-containing 3.
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031536; AAH31536.1; -.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; bromodomain; 2.
DR SMART; SM00297; BRMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS0014; BROMODOMAIN_2; 2.
SQ SEQUENCE 726 AA; 79791 MW; 45C2FC0ACA41DECD CRC64;

Query Match 89.6%; Score 542; DB 11; Length 726;
Best Local Similarity 87.4%; Pred. No. 8.2e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHVKKVVKALWKHGFQFAPFPQVDAVKLGLPDYHKIKQPMDMGTIKRL 60
DB 31 KPGRTNQLQYMNQVNVKTLWKHGFQFAPFPQVDAIKLNPDYHKIKQPMDMGTIKRL 90
QY 61 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 111
DB 90 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQALEKIFLQKVAQMP 140

Query Match 89.6%; Score 542; DB 11; Length 726;
Best Local Similarity 87.4%; Pred. No. 8.2e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHVKKVVKALWKHGFQFAPFPQVDAVKLGLPDYHKIKQPMDMGTIKRL 60
DB 30 KPGRTNQLQYMNQVNVKTLWKHGFQFAPFPQVDAIKLNPDYHKIKQPMDMGTIKRL 89
QY 61 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 111
DB 90 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQALEKIFLQKVAQMP 140

RESULT 10
Q9J125 PRELIMINARY; PRT; 726 AA.
AC Q9J125;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Bromodomain-containing FSH-like protein FSRG2.
GN BRD3 OR 2410084F24RIK OR FSRG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Shang E., Wolgemuth D.J.;
RT "Cloning and expression pattern of FsrG2, a putative murine
RT bromodomain-containing homolog of the Drosophila gene female sterile
RT homeotic".
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF269193; AAF78072.1; -.
DR HSSP; 092831; 1B91.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BRMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS0014; BROMODOMAIN_2; 2.
SQ SEQUENCE 726 AA; 79745 MW; 7AB3B4DAD38A78F4 CRC64;

Query Match 89.6%; Score 542; DB 11; Length 726;
Best Local Similarity 87.4%; Pred. No. 8.2e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHVKKVVKALWKHGFQFAPFPQVDAVKLGLPDYHKIKQPMDMGTIKRL 60
DB 30 KPGRTNQLQYMNQVNVKTLWKHGFQFAPFPQVDAIKLNPDYHKIKQPMDMGTIKRL 89
QY 61 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 111
DB 90 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQALEKIFLQKVAQMP 140

RESULT 11
Q7SXMS PRELIMINARY; PRT; 499 AA.
AC Q7SXMS;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]_TaxID=7955;
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RX MEDLINE=20388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
```

```
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carman R.P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards L., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC055533; AAH55533.1; -.
KW Hypothetical protein.
FT NON_TER 499 499
SQ SEQUENCE 499 AA; 55075 MW; 801BB3D64F9E7216 CRC64;

Query Match 87.6%; Score 530; DB 13; Length 499;
Best Local Similarity 84.7%; Pred. No. 1.2e-50;
Matches 94; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 KGRVTNQLQYLHKVYKMKALWKQFAPFPQPVDAVKLGDPYHKIIPQMDMGTIKRL 60
DB 27 KPGKRTNQLQYMNQVVKTLWKHQFAPFPYTPVDAIKLNPDPYHKVKNPMDMGTIKRL 86

QY 61 ENNYVWAASECMQDNTMTNFTNCYIYNKPTDDIVLMAQLEKIFLOKVASMP 111
DB 87 ENNYVWAASECMQDNTMTNFTNCYIYNKPTDDIVLMAQLEKIFLOKVASMP 137

RESULT 12
Q7SXL8 PRELIMINARY; PRT; 515 AA.
AC Q7SXL8
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE Hypothetical protein (Fragment).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carman R.D., Mullaly S.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Gunaratne P.H.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gay L.J., Hulyk S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
```

```
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC055543; AAH55543.1; -.
KW Hypothetical protein.
FT NON_TER 515 515
SQ SEQUENCE 515 AA; 56253 MW; C3189F567363B32C CRC64;

Query Match 86.6%; Score 524; DB 13; Length 515;
Best Local Similarity 86.5%; Pred. No. 5.7e-50;
Matches 96; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 KGRVTNQLQYLHKVYKMKALWKQFAPFPQPVDAVKLGDPYHKIIPQMDMGTIKRL 60
DB 27 KPGKRTNQLQYMNQVVKTLWKHQFAPFPYTPVDAIKLNPDPYHKIIPQMDMGTIKRL 86

QY 61 ENNYVWAASECMQDNTMTNFTNCYIYNKPTDDIVLMAQLEKIFLOKVASMP 111
DB 87 ENNYVWAASECMQDNTMTNFTNCYIYNKPTDDIVLMAQLEKIFLOKVASMP 137

RESULT 13
Q8QFT7 PRELIMINARY; PRT; 701 AA.
AC Q8QFT7
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21845882; PubMed=11856876;
RA Buchireb N., Grutzner F., Haaf T., Stephens R.J., Elgar G.,
RA Green A.J., Clark M.S.;
RL "Comparative mapping of the human 9q34 region in Fugu rubripes.";
RT Cytochrome. Cell Genet. 94:173-179(2001).
DR EMBL: AJ311635; CAC84085.1; -.
DR InterPro: IPR001487; Bromodomain.
DR Pfam: PF00439; bromodomain; 2.
DR PRINTS: PR00503; BROMODOMAIN.
DR SMART: SM00297; BROMO; 2.
DR PROSITE: PS00633; BROMODOMAIN_1; 2.
DR PROSITE: PS00014; BROMODOMAIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 701 AA; 76356 MW; 74E1F328615E9BD0 CRC64;

Query Match 86.1%; Score 521; DB 13; Length 701;
Best Local Similarity 84.7%; Pred. No. 1.7e-49;
Matches 94; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 KGRVTNQLQYLHKVYKMKALWKQFAPFPQPVDAVKLGDPYHKIIPQMDMGTIKRL 60
DB 28 KPGKRTNQLQYMNQVVKTLWKHQFAPFPYTPVDAIKLNPDPYHKIIPQMDMGTIKRL 87

QY 61 ENNYVWAASECMQDNTMTNFTNCYIYNKPTDDIVLMAQLEKIFLOKVASMP 111
DB 88 ENNYVWAASECMQDNTMTNFTNCYIYNKPTDDIVLMAQLEKIFLOKVASMP 138

RESULT 14
Q7SXP6
```

```

RA Houzelstein D., Bullock S.L., Lynch D.E., Grigorieva E.F.,
RA Wilson V.A., Beddington R.S.P.;
RT "Growth and early post implantation defects in mice mutant for the
RT bromodomain-containing protein Brd4.";
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBU; AF461396; AAL67834.1; -.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00833; BROMODOMAIN_1; 1.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
SQ SEQUENCE 723 AA; 80635 MW; 738F2AE5F58A56BC CRC64;

Query Match 82.6%; Score 500; DB 11; Length 723;
Best Local Similarity 80.2%; Pred. No. 4e-47; Mismatches 0; Gaps 0;
Matches 89; Conservative 10;

QY 1 KPGRVTNQLQYLHKVYMKLWKQHPAWFPQPVDAVKGLGPDYHKIIRKQPMDMGTIKREL 60
Dd 55 KPGKRTNQLQYLRLRVVLTWKQHPAWFPQPVDAVKLNLDPYKLIIRKQPMDMGTIKREL 114

QY 61 ENNYTAAASECQDFNTFTNCYINVKPTDDIVLMAQTLKFIKQKVASMP 111
Dd 115 ENNYTAAASECQDFNTFTNCYINVKPGDDIVLMAEALKFLQKINELP 165

```

Search completed: July 9, 2004, 13:07:35
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:05:11 ; Search time 23 Seconds
(without alignments)
249.151 Million cell updates/sec

Title: US-09-784-553c-19
Perfect score: 605
Sequence: 1 KPGRTVNLQYLHKVVKAL.....IVLMAQTLEKIFLQKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6C-COMB.pep:*
6: /cgn2_6/prodata/2/1aa/6D-COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605	100.0	754	4	US-09-392-714-20
2	542	89.6	726	4	US-09-392-714-21
3	500	82.6	722	4	US-09-392-714-22
4	459	75.9	947	4	US-09-418-780A-1
5	459	75.9	947	4	US-09-392-714-23
6	369	61.0	65	1	US-08-227-536-5
7	369	61.0	65	5	PCT-US95-04682-5
8	204	33.7	65	1	US-08-227-536-6
9	204	33.7	65	5	PCT-US95-04682-6
10	192	31.7	238	4	US-09-257-179-80
11	184.5	30.5	2441	1	US-08-194-468-2
12	184.5	30.5	2441	3	US-08-961-739-2
13	184.5	30.5	2441	4	US-09-514-247A-8
14	184.5	30.5	2441	4	US-09-686-316-2
15	181.5	30.0	2414	1	US-08-227-536-2
16	181.5	30.0	2414	5	PCT-US95-04682-2
17	180.5	29.8	2442	4	US-09-514-247A-10
18	158	26.1	61	4	US-09-418-710-55
19	153	25.3	61	4	US-09-418-710-42
20	149	24.6	1972	4	US-09-418-710-21
21	142	23.5	59	4	US-09-418-710-53
22	139.5	23.1	1525	4	US-09-418-710-69
23	139.5	23.1	1527	4	US-09-418-710-27
24	139.5	23.1	1531	4	US-09-418-710-29
25	139	23.0	1969	4	US-09-418-710-72
26	136.5	22.6	1673	4	US-09-418-710-70
27	136.5	22.6	1674	4	US-09-418-710-1

28	136	22.5	59	4	US-09-418-710-41	Sequence 41, Appl
29	136	22.5	1872	1	US-08-188-582-14	Sequence 14, Appl
30	136	22.5	1872	1	US-08-546-715-14	Sequence 14, Appl
31	136	22.5	1893	1	US-08-188-582-11	Sequence 11, Appl
32	136	22.5	1893	1	US-08-546-715-11	Sequence 11, Appl
33	130	21.5	1878	4	US-09-418-710-13	Sequence 13, Appl
34	129	21.3	65	1	US-08-227-536-4	Sequence 4, Appl
35	129	21.3	65	5	PCT-US95-04682-4	Sequence 4, Appl
36	129	21.3	196	4	US-09-690-454-94	Sequence 94, Appl
37	129	21.3	1876	4	US-09-418-710-71	Sequence 71, Appl
38	126	20.8	59	4	US-09-418-710-38	Sequence 38, Appl
39	126	20.8	59	4	US-09-418-710-50	Sequence 50, Appl
40	123	20.3	589	3	US-08-942-008-2	Sequence 2, Appl
41	120	19.7	59	4	US-09-418-710-51	Sequence 51, Appl
42	119	19.7	59	4	US-09-418-710-39	Sequence 39, Appl
43	110	18.2	59	4	US-09-418-710-52	Sequence 52, Appl
44	110	18.2	65	1	US-08-227-536-8	Sequence 8, Appl
45	110	18.2	65	5	PCT-US95-04682-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-392-714-20
; Sequence 20, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O. Barbara
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE OF INVENTION: Therefor
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-20

Query Match Similarity 100.0%; Score 605; DB 4; Length 754;
Best Local Similarity 100.0%; Pred. No. 6.5e-71;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KPGRTVNLQYLHKVVKALWKHQFAWPFQPDVAVKGLDYLHKLKQPMNDGTIKRRL	60
DB	24	KPGRTVNLQYLHKVVKALWKHQFAWPFQPDVAVKGLDYLHKLKQPMNDGTIKRRL	83
QY	61	ENNYWAASECQDPNTMTNCYTNKPTDDIVLMAQTLEKIFLQKVASMP	111
DB	84	ENNYWAASECQDPNTMTNCYTNKPTDDIVLMAQTLEKIFLQKVASMP	134

RESULT 2
US-09-392-714-21
; Sequence 21, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE OF INVENTION: Therefor

FILE REFERENCE: L0461/7062
CURRENT APPLICATION NUMBER: US/09/392,714A
CURRENT FILING DATE: 1999-09-09
EARLIER APPLICATION NUMBER: PCT/US98/14679
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 21
LENGTH: 726
TYPE: PRT
ORGANISM: Homo sapiens
US-09-392-714-21

Query Match 89.6%; Score 542; DB 4; Length 726;
Best Local Similarity 87.4%; Pred. No. 1.2e-62;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 KPGRTNQLQYLHKVVKALWKHOFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60
Db 31 KPGRTNQLQYVNVVVKLWKHOFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 90
QY 61 ENNYWASACQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 91 ENNYWASACQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 141

RESULT 3
US-09-392-714-22
Sequence 22, Application US/09392714A
Patent No. 6686147
GENERAL INFORMATION:
APPLICANT: Scanlan, Matthew J.
APPLICANT: Gure, Ali O.
APPLICANT: Williams, Barbara
APPLICANT: Chen, Yao-Tseng
APPLICANT: Old, Lloyd J.
TITLE OF INVENTION: Cancer Associated Antigens and Uses
TITLE OF INVENTION: Therefor
FILE REFERENCE: L0461/7062
CURRENT APPLICATION NUMBER: US/09/392,714A
CURRENT FILING DATE: 1999-09-09
EARLIER APPLICATION NUMBER: PCT/US98/14679
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 22
LENGTH: 722
TYPE: PRT
ORGANISM: Homo sapiens
US-09-392-714-22

Query Match 82.6%; Score 500; DB 4; Length 722;
Best Local Similarity 80.2%; Pred. No. 4.2e-57;
Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
QY 1 KPGRTNQLQYLHKVVKALWKHOFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60
Db 55 KPGRTNQLQYLVRVLKTLWKHOFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 114
QY 61 ENNYWASACQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 115 ENNYWASACQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 165

RESULT 4
US-09-418-780A-1
Sequence 1, Application US/09418780A
Patent No. 6504009
GENERAL INFORMATION:
APPLICANT: Jores, Michael H.
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
FILE REFERENCE: 06501-043001
CURRENT APPLICATION NUMBER: US/09/418,780A

CURRENT FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: PCT/JP98/01782
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: JAPAN 9/116402
PRIOR FILING DATE: 1997-04-18
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 947
TYPE: PRT
ORGANISM: Homo sapiens
US-09-418-780A-1

Query Match 75.9%; Score 459; DB 4; Length 947;
Best Local Similarity 73.0%; Pred. No. 1.5e-51;
Matches 81; Conservative 16; Mismatches 14; Indels 0; Gaps 0;
QY 1 KPGRTNQLQYLHKVVKALWKHOFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60
Db 24 KNGRLTNQLQYLQKVLKDLWKHSFSPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 83
QY 61 ENNYWASACQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 84 ENNYWASACQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 134

RESULT 5
US-09-392-714-23
Sequence 23, Application US/09392714A
Patent No. 6686147
GENERAL INFORMATION:
APPLICANT: Scanlan, Matthew J.
APPLICANT: Gure, Ali O.
APPLICANT: Williams, Barbara
APPLICANT: Chen, Yao-Tseng
APPLICANT: Old, Lloyd J.
TITLE OF INVENTION: Cancer Associated Antigens and Uses
TITLE OF INVENTION: Therefor
FILE REFERENCE: L0461/7062
CURRENT APPLICATION NUMBER: US/09/392,714A
CURRENT FILING DATE: 1999-09-09
EARLIER APPLICATION NUMBER: PCT/US98/14679
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 947
TYPE: PRT
ORGANISM: Homo sapiens
US-09-392-714-23

Query Match 75.9%; Score 459; DB 4; Length 947;
Best Local Similarity 73.0%; Pred. No. 1.5e-51;
Matches 81; Conservative 16; Mismatches 14; Indels 0; Gaps 0;
QY 1 KPGRTNQLQYLHKVVKALWKHOFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60
Db 24 KNGRLTNQLQYLQKVLKDLWKHSFSPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 83
QY 61 ENNYWASACQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 84 ENNYWASACQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 134

RESULT 6
US-08-227-536-5
Sequence 5, Application US/08227536
Patent No. 5658784
GENERAL INFORMATION:
APPLICANT: Eckner, Richard
APPLICANT: Ewen, Mark
APPLICANT: Livingston, David
TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION

TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: DFCI-308XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-027-536-5

Query Match 61.0%; Score 369; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.3e-41;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0

[illegible]

```

1  RESULT 7
2  PCT-US95-04682-5
3  ; Sequence 5, Application PC/TUS9504682
4  ; GENERAL INFORMATION:
5  ; APPLICANT:
6  ; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
7  ; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
8  ; NUMBER OF SEQUENCES: 13
9  ; CORRESPONDENCE ADDRESS:
10 ; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
11 ; STREET: Post Office Square
12 ; CITY: Boston
13 ; STATE: MA
14 ; COUNTRY: US
15 ; ZIP: 02109
16 ; COMPUTER READABLE FORM:
17 ; MEDIUM TYPE: Floppy disk
18 ; COMPUTER: IBM PC compatible
19 ; OPERATING SYSTEM: PC-DOS/MS-DOS
20 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
21 ; CURRENT APPLICATION DATA:
22 ; APPLICATION NUMBER: PCT/US95/04682
23 ; FILING DATE:
24 ; CLASSIFICATION:

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-April-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-306XG999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
CT-0595-04682-5

Query Match 61.0%; Score 369; DB 5; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.3e-41;
Matches 65; Conservative 0; Mismatches 0; Indels

Qy	24	QAFNFRCPVDVAVLGLPDYHKIIKQPMDMGTIIKRLENNYVWAASECMQDFNMTNCTCY	83
Db	1	QAFNFRCPVDVAVLGLPDYHKIIKQPMDMGTIIKRLENNYVWAASECMQDFNMTNCTCY	60
Qy	34	IYNKP	88
Db	61	IYNKP	65

RESULT 8
US-08-227-536-6

Patent No. 5656784
GENERAL INFORMATION:
APPLICANT: Eckner, Richard
APPLICANT: Ewen, Mark
APPLICANT: Livingston, David
TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square

CITY: BOSTON
 STATE: MA
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/227,536
 FILING DATE: 14-APR-1994
 CLASSIFICATION: 436
 ATTORNEY/AGENT INFORMATION:
 NAME: Williams Ph.D., Kathleen A.
 REGISTRATION NUMBER: 34,380
 REFERENCE/DOCKET NUMBER: DFCI-308XX
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-2290
 TELEFAX: (617) 451-0313
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 65 amino acids
 TYPE: amino acid

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
US-08-227-536-6

Query Match
Best Local Similarity 33.7%; Score 204; DB 1; Length 65;
Matches 37; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 25 FAWPFRQPDVAVKGLPDYHKIKQPMDMGTIKRLENNYWAASECMQDFTMTNCYI 84
Db 2 YAWPFYKPDASALGHDYHDIKHPMDLSTVKRKMENRDYRDAQEFADVRLMFSNCYK 61

QY 85 YNKP 88
Db 62 YNPP 65

RESULT 9
PCT-US95-04682-6
; Sequence 6, Application PC/TUS9504682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04682
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holliday C. Heine, Ph.D.
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: DFCI-308Xg999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 65 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
PCT-US95-04682-6

Query Match
Best Local Similarity 33.7%; Score 204; DB 5; Length 65;
Matches 37; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 25 FAWPFRQPDVAVKGLPDYHKIKQPMDMGTIKRLENNYWAASECMQDFTMTNCYI 84
Db 2 YAWPFYKPDASALGHDYHDIKHPMDLSTVKRKMENRDYRDAQEFADVRLMFSNCYK 61

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
US-09-257-179-80

Query Match
Best Local Similarity 31.7%; Score 192; DB 4; Length 238;
Matches 36; Conservative 19; Mismatches 37; Indels 2; Gaps 1;

QY 14 KVMKALWKHOFANFPFQPDVAVKGLPDYHKIKQPMDMGTIKRLENNYWAASECMQ 73
Db 129 KRVLSLQAHXWAMPFLEPVD--PNDAPDYGVKPEMDLATMEERVQRRYYEKLTFVA 186

QY 74 DFTMTFTNCYIYNKPTDDVLMAQTLKIFLOKV 107
Db 187 DWTKIFDNCRYNPSDSFFFOCAVLSFFVQKL 220

RESULT 11
US-08-194-468-2
; Sequence 2, Application US/08194468
; Patent No. 5750336
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
; TITLE OF INVENTION: RESPONSIVE GENES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,468
; FILING DATE: 10-FEB-1994
; CLASSIFICATION: 435
```

Fri Jul 9 16:10:21 2004

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9672
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)-546-4737
; TELEFAX: (619)-546-9392
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-194-468-2

Query Match          30.5%; Score 184.5; DB 1; Length 2441;
Best Local Similarity 38.0%; Pred. No. 7.8e-15;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAWPFRQPVDAVKGLDPYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75
Db 1099 LEALYRQDPESLPFRQPVDPQLLGIPDYFDIVKNPMDLSTIKRKLDTGQYQEPWQYVDDV 1158

QY 76 NTWFTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
Db 1159 RLMFNNAWLYNKRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 12
US-08-961-739-2
; Sequence 2, Application US/08961739A
; Patent No. 6063583
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
; FILE REFERENCE: SALK1650-1
; CURRENT APPLICATION NUMBER: US/08/961,739A
; CURRENT FILING DATE: 1997-10-31
; EARLIER APPLICATION NUMBER: US 194,468
; EARLIER FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: Mus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2441)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-08-961-739-2

Query Match          30.5%; Score 184.5; DB 3; Length 2441;
Best Local Similarity 38.0%; Pred. No. 7.8e-15;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAWPFRQPVDAVKGLDPYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75
Db 1099 LEALYRQDPESLPFRQPVDPQLLGIPDYFDIVKNPMDLSTIKRKLDTGQYQEPWQYVDDV 1158

QY 76 NTWFTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
Db 1159 RLMFNNAWLYNKRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 13
US-09-514-247A-8
; Sequence 8, Application US/09514247A
; Patent No. 6355361
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko

; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPAI
; FILE REFERENCE: TANIGUCHI=6
; CURRENT APPLICATION NUMBER: US/09/514,247A
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: mouse
; US-09-514-247A-8

Query Match          30.5%; Score 184.5; DB 4; Length 2441;
Best Local Similarity 38.0%; Pred. No. 7.8e-15;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAWPFRQPVDAVKGLDPYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75
Db 1099 LEALYRQDPESLPFRQPVDPQLLGIPDYFDIVKNPMDLSTIKRKLDTGQYQEPWQYVDDV 1158

QY 76 NTWFTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
Db 1159 RLMFNNAWLYNKRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 14
US-09-686-316-2
; Sequence 2, Application US/09686316
; Patent No. 6646115
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
; FILE REFERENCE: SALK1650-1
; CURRENT APPLICATION NUMBER: US/09/686,316
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US/08/961,739
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: US 194,468
; PRIOR FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: Mus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2441)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-686-316-2

Query Match          30.5%; Score 184.5; DB 4; Length 2441;
Best Local Similarity 38.0%; Pred. No. 7.8e-15;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAWPFRQPVDAVKGLDPYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75
Db 1099 LEALYRQDPESLPFRQPVDPQLLGIPDYFDIVKNPMDLSTIKRKLDTGQYQEPWQYVDDV 1158

QY 76 NTWFTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
Db 1159 RLMFNNAWLYNKRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 15
US-08-227-536-2
; Sequence 2, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
```

APPLICANT: Eckner, Richard
APPLICANT: Ewen, Mark
APPLICANT: Livingston, David
TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: DFCI-308XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-536-2

Query Match 30.0%; Score 181.5; DB 1; Length 2414;
Best Local Similarity 38.0%; Pred. No. 1.9e-14;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;
QY 17 MKALWKHQ-FAMPRQPVDAVKLGDPDYHKIIKQPMDMGTIKRRLENNYYWAAASECMQDF 75
DB 1062 LEALYRQDPESLPFRQPVDFQIGIPDYFDIVKSPMDLSTIKRKLDGTQYQEPWQYVDDI 1121
QY 76 NTMETNCYVYNKPTDDIVLNAQTLEKIFLQKV 107
DB 1122 WLFNNALVNRKTSRVKYCKSLSEVFQEI 1153

Search completed: July 9, 2004, 13:08:43
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:07:41 ; Search time 47 Seconds
(without alignments)
736.658 Million cell updates/sec

Title: US-09-784-553C-19
Perfect score: 605
Sequence: 1 KGRVTNQLQYLHKVVMKAL.....IVLMAQTLEKIFLOKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues
Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	605	100.0	111	12	US-09-784-553C-19
2	605	100.0	111	15	US-10-209-201C-19
3	605	100.0	801	14	US-10-146-473-42
4	542	89.6	111	12	US-09-784-553C-20
5	542	89.6	111	15	US-10-209-201C-20
6	542	89.6	140	9	US-09-764-864-1572
7	542	89.6	235	9	US-09-764-864-1159
8	500	82.6	1390	12	US-10-092-900A-224
9	498	82.3	111	12	US-09-784-553C-21
10	498	82.3	111	15	US-10-209-201C-21
11	459	75.9	947	14	US-10-293-822-1
12	269	44.5	113	12	US-09-784-553C-23
13	269	44.5	113	15	US-10-209-201C-23
14	256.5	42.4	113	12	US-09-784-553C-25
15	256.5	42.4	113	15	US-10-209-201C-25

16	244	40.3	113	12	US-09-784-553C-24	Sequence 24, Appl
17	244	40.3	113	15	US-10-209-201C-24	Sequence 24, Appl
18	224	37.0	268	12	US-10-424-599-219965	Sequence 219965, A
19	223	36.9	577	12	US-10-425-114-37396	Sequence 37396, A
20	215	35.5	578	15	US-10-369-493-22750	Sequence 22750, A
21	208.5	34.5	128	12	US-10-424-599-244688	Sequence 244688, A
22	204	33.7	386	14	US-10-176-884-13	Sequence 13, Appl
23	204	33.7	386	14	US-10-177-478-4	Sequence 4, Appl
24	199.5	33.0	656	16	US-10-437-963-154093	Sequence 154093, A
25	198	32.7	110	12	US-09-784-553C-7	Sequence 7, Appl
26	198	32.7	110	15	US-10-209-201C-7	Sequence 7, Appl
27	196	32.4	340	12	US-10-425-114-41650	Sequence 41650, A
28	196	32.4	610	12	US-10-424-599-198828	Sequence 198828, A
29	195	32.2	128	9	US-09-764-864-1571	Sequence 1571, Ap
30	195	32.2	167	16	US-10-437-963-116598	Sequence 116598, A
31	195	32.2	791	16	US-10-437-963-172902	Sequence 172902, A
32	192	31.7	237	12	US-10-220-120-243	Sequence 243, App
33	192	31.7	238	9	US-09-729-835-80	Sequence 80, Appl
34	192	31.7	238	16	US-10-373-809-80	Sequence 80, Appl
35	192	31.7	2781	12	US-10-263-929-122	Sequence 122, App
36	190	31.4	273	12	US-10-424-599-226865	Sequence 226865, A
37	190	31.4	346	12	US-10-425-114-64125	Sequence 64125, A
38	189.5	31.3	563	16	US-10-437-963-198539	Sequence 198539, A
39	189	31.2	372	12	US-10-425-114-38433	Sequence 38433, A
40	188	31.1	751	16	US-10-437-963-172901	Sequence 172901, A
41	187	30.9	516	16	US-10-437-963-116164	Sequence 116164, A
42	187	30.9	832	12	US-09-784-553C-2	Sequence 2, Appl
43	187	30.9	832	15	US-10-209-201C-2	Sequence 2, Appl
44	184.5	30.5	112	12	US-09-784-553C-13	Sequence 13, Appl
45	184.5	30.5	112	15	US-10-209-201C-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-784-553C-19
; Sequence 19, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: ZHOU, MING-MING
; APPLICANT: AGGARWAL, ANEEL
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784, 553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; TYPE: PRT
; LENGTH: 111
; ORGANISM: Homo sapiens
US-09-784-553C-19

Query Match	100.0%	Score 605	DB 12	Length 111
Best Local Similarity	100.0%	Pred. No. 5.6e-66		
Matches 111	Conservative	0	Mismatches 0	Indels 0
Gaps	0			
Qy	1	KGRVTNQLQYLHKVVMKALWKHQFAWPPQVDAVKLGLPDYHKKIKQPMDMGTIKRRL	60	
Db	1	KGRVTNQLQYLHKVVMKALWKHQFAWPPQVDAVKLGLPDYHKKIKQPMDMGTIKRRL	60	
Qy	61	ENNYWAASECMDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP	111	
Db	61	ENNYWAASECMDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP	111	
RESULT 2				
US-10-209-201C-19				
; Sequence 19, Application US/10209201C				
; Publication No. US20040009613A1				

```
RESULT 3
US-10-146-473-42
; Sequence 42, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-42

Query Match      100.0%; Score 605; DB 14; Length 801;
Best Local Similarity 100.0%; Pred. No. 6.3e-65;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
DB 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
QY 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111

RESULT 4
US-10-146-473-42
; Sequence 42, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-42

Query Match      100.0%; Score 605; DB 14; Length 801;
Best Local Similarity 100.0%; Pred. No. 6.3e-65;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
DB 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
QY 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
```

```
RESULT 4
US-09-784-553C-20
; Sequence 20, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: ZHOU, MING-MING
; APPLICANT: AGGARWAL, ANEEL
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784,553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-553C-20

Query Match      89.6%; Score 542; DB 12; Length 111;
Best Local Similarity 87.4%; Pred. No. 2.9e-58;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
DB 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
QY 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111

RESULT 5
US-10-209-201C-20
; Sequence 20, Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-201C-20

Query Match      89.6%; Score 542; DB 15; Length 111;
Best Local Similarity 87.4%; Pred. No. 2.9e-58;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
DB 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
QY 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
```

```
RESULT 5
US-10-209-201C-20
; Sequence 20, Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-201C-20

Query Match      89.6%; Score 542; DB 15; Length 111;
Best Local Similarity 87.4%; Pred. No. 2.9e-58;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
DB 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPVDVAVKLGLEDYHKKIKQPMDMGTIKRRL 60
QY 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 61 ENNYWAASECMQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
```

Query Match	82.6%;	Score 500;	DB 12;	Length 1390;
Best Local Similarity	80.2%;	Pred. No. 8.9e-52;		
Matches 89:	Conservative	10;	Mismatches 12;	Indels 0

[illegible]

```

RESULT 9
US-09-784-553C-21
; Sequence 21, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: ZHOU, MING-MING
; APPLICANT: AGGARWAL, ANEEL
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784,553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-784-553C-21

```

```

RESULT 10
US-10-209-201C-21
; Sequence 21, Application US/10209201C
; Publication No. US2004009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Agarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-209-201C-21

```

[illegible]

```

RESULT 11
US-10-293-822-1
; Sequence 1, Application US/10293822
; Publication No. US20030083470A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/10/293,822
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US/09/418,780
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JF98/01782
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JAPAN 9/116402
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-822-1

```

```

RESULT 12
US-09-784-553C-23
; Sequence 23, Application US/09784553C
; Publication NO. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: ZHOU, MING-MING
; APPLICANT: AGGARWAL, ANEEL
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784,553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-553C-23

```

```
Query Match      44.5%; Score 269; DB 13; Length 113;
Best Local Similarity 46.0%; Pred.No. 7.8e-25;
Matches 52; Conservative 21; Mismatches 38; Indels 2; Gaps 1

QY      1 KPRVTNQLVYLHKVVKALWKHO--FAWPRQPVDVKLGLPDYHKLIIKPMDMTIKR 58
          ||| | : : : : : : : : : : : : : : : : : : : : : : : :
```



```

Db      1 KGGKLSQKHNGIILKELLSSKKHAAAYAFYKPVDASALGLHDVHDIIKHPMDLSTVKR   60
Qy      59 RLENNYYAASECMQDFTMTFCNTCYINYKNPTDDIVLMAQTLEKIIFLOKVASMP    111
        :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     61 KMENRDYDAQEFAADVRLMFNSCNYKNPPDHVVAMARKLDQVFERYAKMP    113

RESULT 13
US-10-209-201C-23
; Sequence 23, Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Regarwal, Ansel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-201C-23

Query Match          44.5%; Score 269; DB 15; Length 113;
Best Local Similarity 46.0%; Pred.No. 7.se-25;
Matches 52; Conservative 21; Mismatches 38; Indels 2; Gaps

Qy      1 KPGRFYNQLYLHKVVMKALKWQHQQ--FWAPFRQPVDVAVLGPLDYHKLIKQMMDMGTIKR   58
        :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     61 KGGKLSQKHNGIILKELLSSKKHAAAYAFYKPVDASALGLHDVHDIIKHPMDLSTVKR   60

Qy      59 RLENNYYAASECMQDFTMTFCNTCYINYKNPTDDIVLMAQTLEKIIFLOKVASMP    111
        :|::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     61 KMENRDYDAQEFAADVRLMFNSCNYKNPPDHVVAMARKLDQVFERYAKMP    113

```

[illegible][illegible]

Search completed: July 9, 2004, 13:13:32
Job time : 48 secs

